



TESIS - SS14 2501

**PEMODELAN STATUS BEKERJA IBU RUMAH TANGGA
MENGUNAKAN MODEL MULTILEVEL
DENGAN RESPON BINER**

FANNY AYU OCTAVIANA

NRP. 1315 201 029

DOSEN PEMBIMBING :

Dr. Wahyu Wibowo, M.Si

Dr. R. Mohamad Atok, M.Si

PROGRAM MAGISTER

JURUSAN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

INSTITUT TEKNOLOGI SEPULUH NOPEMBER

SURABAYA

2017



THESIS - SS14 2501

MODELLING WORK STATUS OF HOUSEWIFE USING MULTILEVEL MODEL FOR BINARY RESPONSE

FANNY AYU OCTAVIANA

NRP. 1315 201 029

SUPERVISOR :

Dr. Wahyu Wibowo, M.Si

Dr. R. Mohamad Atok, M.Si

PROGRAM OF MAGISTER

DEPARTEMENT OF STATISTICS

FACULTY OF MATHEMATICS AND NATURAL SCIENCES

INSTITUT TEKNOLOGI SEPULUH NOPEMBER

SURABAYA

2017


**PEMODELAN STATUS BEKERJA IBU RUMAH TANGGA
MENGUNAKAN MODEL MULTILEVEL
DENGAN RESPON BINER**

**Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Sains (M.Si)
di
Institut Teknologi Sepuluh Nopember
Oleh :**

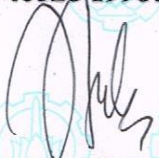
**FANNY AYU OCTAVIANA
NRP. 1315 201 029**

**Tanggal Ujian : 6 Januari 2017
Periode Wisuda : Maret 2017**


Disetujui oleh :


**1. Dr. Wahyu Wibowo, M.Si
NIP. 19740328 199802 1 001**

(Pembimbing I)


**2. Dr. R. Mohamad Atok, M.Si
NIP. 19710915 199702 1 001**

(Pembimbing II)


**3. Dr. Drs. Agus Suharsono, M.S
NIP. 19580823 198403 1 003**


(Penguji I)


**4. Dr. rer. pol. Heri Kuswanto, M.Si
NIP. 19820326 200312 1 004**

(Penguji II)

**an. Direktur Program Pascasarjana
Asisten Direktur**

Direktur Program Pascasarjana


**Prof. Dr. Ir. H. Widjaya, M.Eng.
NIP. 19611021 198603 1 001**

**Prof. Ir. Djauhar Manfaat, M.Sc, Ph.D
NIP. 19601202 198701 1 001**

PEMODELAN STATUS BEKERJA IBU RUMAH TANGGA MENGUNAKAN MODEL MULTILEVEL DENGAN RESPON BINER

Nama Mahasiswa : Fanny Ayu Octaviana
NRP : 1315 201 029
Jurusan : Statistika FMIPA ITS
Pembimbing : Dr. Wahyu Wibowo, M.Si
Co-Pembimbing : Dr. R. Mohamad Atok. M.Si

ABSTRAK

Kurangnya pendapatan yang dihasilkan oleh suami sebagai kepala rumah tangga dan pencari nafkah membuat sebagian besar wanita ikut serta bekerja guna memenuhi kebutuhan keluarga. Dalam penelitian ini, akan dilakukan pemodelan multilevel dengan respon biner pada status bekerja ibu rumah tangga berdasarkan pendidikan terakhir, usia dan pengeluaran rumah tangga. Penelitian ini bertujuan untuk mengetahui hubungan antara status bekerja ibu rumah tangga berdasarkan variabel-variabel pada setiap level dan model yang sesuai dari penelitian tentang status bekerja ibu rumah tangga. Data yang digunakan adalah data sekunder yang diperoleh dari *Indonesia Family Life Survey* (IFLS) tahun 2014, yaitu total keseluruhan data adalah 4431 wanita yang memiliki status menikah atau pernah menikah dan sebanyak 99 data rata-rata gaji tenaga kerja wanita yang bersesuaian dengan data dalam Kecamatan. Model multilevel yang digunakan dalam penelitian ini adalah model regresi 2 level dengan respon biner. Variabel penelitian dalam penelitian ini adalah status bekerja ibu rumah tangga sebagai variabel dependen (Y), dengan dua kategori yaitu 0 untuk wanita tidak bekerja dan 1 untuk wanita bekerja, serta pendidikan terakhir, usia dan pengeluaran rumah tangga sebagai variabel independen (X) pada level 1 (individu) dan rata-rata gaji tenaga kerja wanita sebagai variabel independen (Z) pada level 2 (Kecamatan). Langkah pertama dalam menganalisis data dengan pemodelan multilevel adalah melakukan pemodelan menggunakan regresi logistik biner pada tingkat 1 (individu), kemudian melakukan pemodelan menggunakan model regresi 2-level dengan respon biner pada level gabungan yaitu level 1 dan level 2 (Kecamatan). Dari hasil analisis dapat diperoleh kesimpulan bahwa model yang terbaik untuk kasus status bekerja ibu rumah tangga di Indonesia adalah menggunakan model regresi 2-level dengan respon biner yang mengikutsertakan variabel rata-rata gaji tenaga kerja wanita karena nilai *Deviance* yang dihasilkan paling kecil dibandingkan dengan model lainnya. Artinya, keputusan ibu rumah tangga untuk ikut serta dalam dunia kerja dipengaruhi oleh variabel pendidikan terakhir dan variabel pengeluaran rumah tangga (level 1/individu), serta dipengaruhi oleh adanya daerah tempat tinggal (level 2/Kecamatan).

Kata kunci : *Indonesia Family Life Survey*, Model Multilevel, Respon Biner, Status Bekerja Ibu Rumah Tangga.

MODELLING WORK STATUS OF HOUSEWIFE USING MULTILEVEL MODEL FOR BINARY RESPONSE

Name : Fanny Ayu Octaviana
NRP : 1315 201 029
Departement : Statistika FMIPA ITS
Supervisor : Dr. Wahyu Wibowo, M.Si
Co-Supervisor : Dr. R. Mohamad Atok. M.Si

ABSTRACT

The lack of revenue generated by husband as head of household and breadwinner make most women opt to work in order to meet needs of the family. In this study, multilevel modeling will be conducted on work status of women based on the latest education, age and household expenditure. This study aims to determine the relationship between work status of women based on variables at each level and an appropriate model of research on work status of women. The data used is secondary data obtained from the Indonesia Family Life Survey (IFLS) in 2014, i.e. the total of data is 4431 women who have ever been married or married status and as many as 99 data is the average salary of women that corresponds to data in the sub-district. Multilevel model used in this study is a 2-level regression model for binary response. The variables of this research is work status of women as the response variable (Y), with two categories: 0 to women did not work and 1 to women work, and the latest education, age and household expenditure as a predictor variable (X) at level 1 (individual) and the average salary of female workers as a predictor variable (Z) at level 2 (sub-district). The first step in analyzing the data with multilevel modeling is modeling with binary logistic regression at level 1 (individual), then do the modeler with 2-level regression model for binary response to determine the effects of their level 2 (sub-district). From the analysis, it can be concluded that the best model for work status of housewife case in Indonesia are using a 2-level regression model for binary response with the latest education and household expenditure variables that influence work status of housewife because the generated of deviance value is most small compared to other models. That is, the housewife's decision to participate in the work world is influenced by the latest education and household expenditure variables (level 1/individual), and also influenced by the area of residence (level 2/sub-district).

Keywords : Indonesia Family Life Survey, Multilevel Model, Binary Logistic Regression, Binary Response, Work Status of Housewife.

KATA PENGANTAR

Puji syukur kepada Allah SWT atas rahmat, hidayah, serta karunia-Nya sehingga penulis dapat menyelesaikan penyusunan Tesis ini dengan judul **“PEMODELAN STATUS BEKERJA IBU RUMAH TANGGA MENGGUNAKAN MODEL MULTILEVEL DENGAN RESPON BINER”**. Penulisan Tesis ini merupakan pelaksanaan kewajiban akademis sebagai salah satu syarat yang harus dipenuhi untuk menyelesaikan pendidikan Program Magister Statistika pada Program Pascasarjana di Institut Teknologi Sepuluh Nopember Surabaya. Dalam penyelesaian tesis ini, penulis menyadari bahwa tesis ini tidak akan terselesaikan tanpa bantuan dan arahan serta dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih sebesar-besarnya kepada :

1. Bapak Dr. Suhartono, M.Sc, selaku Ketua Jurusan Statistika FMIPA ITS Surabaya.
2. Bapak Dr.rer.pol. Heri Kuswanto, M.Si, selaku Ketua Program Studi Pascasarjana Jurusan Statistika FMIPA ITS Surabaya dan dosen wali yang telah membimbing dan mengarahkan sejak awal perkuliahan.
3. Bapak Dr. Wahyu Wibowo, M.Si dan Bapak Dr. R. Mohamad Atok, M.Si, selaku dosen Pembimbing dan Co-pembimbing Tesis yang telah meluangkan waktu untuk memberikan bimbingan, saran, dan dukungan kepada penulis untuk menyelesaikan Tesis ini.
4. Bapak Dr. Drs. Agus Suharsono, M.S dan Bapak Dr.rer.pol. Heri Kuswanto, M.Si, selaku dosen penguji yang telah memberikan tambahan ilmu selama proses perbaikan untuk Tesis ini.
5. Seluruh Bapak/Ibu Dosen pengajar, staf administrasi, dan staf perpustakaan Jurusan Statistika ITS yang telah berkenan memberikan waktu dan bantuan.
6. Kedua orang tua, kakak dan adikku yang tersayang, atas segala doa, bimbingan, bantuan, motivasi dan kasih sayang yang tak pernah surut dalam menemani setiap langkahku.

7. Teman-teman seperjuangan mahasiswa Magister Statistika ITS Angkatan 2015, atas dukungan, motivasi, dan kebersamaan selama menempuh pendidikan bersama.
8. Semua pihak yang telah membantu dalam penulisan Tesis ini yang tidak dapat penulis sebutkan satu per satu.

Penulis menyadari bahwa masih terdapat kekurangan pada Tesis ini, maka saran dan kritik dari semua pihak yang bersifat membangun sangat penulis harapkan demi kesempurnaan penulisan selanjutnya. Semoga laporan ini dapat memberikan sumbangan yang bermanfaat bagi semua pihak.

Surabaya, Januari 2017

Penulis

DAFTAR ISI

| | |
|--------------------------------------------------------------------------|-------------|
| HALAMAN JUDUL | i |
| LEMBAR PENGESAHAN | ii |
| ABSTRAK | iii |
| ABSTRACT | v |
| KATA PENGANTAR..... | vii |
| DAFTAR ISI..... | ix |
| DAFTAR GAMBAR..... | xi |
| DAFTAR TABEL | xiii |
| DAFTAR LAMPIRAN | xv |
| BAB 1 PENDAHULUAN | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Rumusan Masalah | 5 |
| 1.3 Tujuan Penelitian..... | 5 |
| 1.4 Manfaat Penelitian..... | 5 |
| 1.5 Batasan Masalah..... | 6 |
| BAB 2 LANDASAN TEORI | 7 |
| 2.1 Multikolinearitas | 7 |
| 2.2 Model Regresi..... | 7 |
| 2.3 Model Regresi Multilevel | 8 |
| 2.3.1 Model Regresi 2 Level | 9 |
| 2.3.2 Model Regresi 2 Level Dengan <i>Random Intercept</i> | 12 |
| 2.4 Model Regresi Logistik Biner..... | 14 |
| 2.5 Model Regresi Multilevel Dengan Respon Biner | 15 |
| 2.6 Estimasi Parameter Model Regresi Multilevel Dengan Respon Biner | 16 |
| 2.6.1 <i>Generalized Linear Mixed Model</i> (GLMM)..... | 16 |
| 2.6.2 <i>Maximum Likelihood Estimation</i> (MLE)..... | 16 |
| 2.7 Pengujian Signifikansi Parameter | 19 |
| 2.8 Pemilihan Model Terbaik | 19 |
| 2.9 Ketepatan Klasifikasi | 20 |
| 2.10 Tenaga Kerja Wanita | 21 |

| | |
|----------------------------------------------------------------------------------------------|-----------|
| BAB 3 METODOLOGI PENELITIAN | 25 |
| 3.1 Sumber Data | 25 |
| 3.2 Variabel Penelitian | 25 |
| 3.3 Struktur Data Penelitian | 27 |
| 3.4 Langkah-Langkah Analisis | 27 |
| BAB 4 ANALISIS DAN PEMBAHASAN..... | 33 |
| 4.1 Estimasi Parameter Model Regresi Multilevel Respon Biner | 33 |
| 4.2 Karakteristik Data | 38 |
| 4.2.1 Status Bekerja Ibu Rumah Tangga..... | 38 |
| 4.2.2 Pendidikan Terakhir | 39 |
| 4.2.3 Usia | 40 |
| 4.2.4 Pengeluaran Rumah Tangga | 40 |
| 4.2.5 Rata-rata Gaji Wanita..... | 41 |
| 4.3 Uji Multikolinearitas | 42 |
| 4.4 Analisis Regresi Logistik Biner | 42 |
| 4.4.1 Uji Estimasi Parameter Secara Serentak | 42 |
| 4.4.2 Uji Estimasi Parameter Secara Parsial | 43 |
| 4.5 Analisis Regresi Multilevel Respon Biner untuk <i>Random Intercept</i> | 45 |
| 4.5.1 Analisis Regresi Multilevel Respon Biner Tanpa Variabel Rata-rata Gaji Wanita..... | 45 |
| 4.5.2 Analisis Regresi Multilevel Respon Biner Dengan Variabel Rata-rata Gaji Wanita..... | 47 |
| 4.6 Pemilihan Model Terbaik | 49 |
| 4.7 Ketepatan Klasifikasi | 50 |
| BAB 5 KESIMPULAN DAN SARAN | 53 |
| 5.1 Kesimpulan | 53 |
| 5.2 Saran | 54 |
| DAFTAR PUSTAKA | 55 |
| LAMPIRAN..... | 59 |

DAFTAR GAMBAR

| | | |
|-------------------|-------------------------------------------------------|----|
| Gambar 3.1 | Diagram alir penelitian | 31 |
| Gambar 4.1 | Persentase status bekerja ibu rumah tangga | 38 |
| Gambar 4.2 | Persentase pendidikan terakhir ibu rumah tangga | 39 |

DAFTAR TABEL

| | | |
|-------------------|-------------------------------------------------------------------------------------|----|
| Tabel 2.1 | Struktur Data Model 2 Level..... | 12 |
| Tabel 2.2 | Struktur Data Regresi Logistik Biner..... | 15 |
| Tabel 2.3 | Tabel Ketepatan Klasifikasi Untuk 2 Kategori | 20 |
| Tabel 2.4 | Tabel Sensitivitas Untuk 2 Kategori | 21 |
| Tabel 3.1 | Variabel Penelitian | 26 |
| Tabel 3.2 | Struktur Data Penelitian Model 2 Level..... | 27 |
| Tabel 4.1 | Statistika Deskriptif Untuk Usia | 40 |
| Tabel 4.2 | Statistika Deskriptif Untuk Pengeluaran Rumah Tangga | 41 |
| Tabel 4.3 | Statistika Deskriptif Untuk Rata-rata Gaji Wanita Setiap Kecamatan..... | 41 |
| Tabel 4.4 | Nilai VIF Pada Variabel Independen | 42 |
| Tabel 4.5 | Uji Estimasi Parameter Serentak Analisis Regresi Logistik Biner | 43 |
| Tabel 4.6 | Uji Estimasi Parameter Parsial Analisis Regresi Logistik Biner ... | 44 |
| Tabel 4.7 | Uji Estimasi Parameter Parsial Analisis Regresi Logistik Tanpa X_2 | 45 |
| Tabel 4.8 | Uji Estimasi Parameter Analisis Regresi Multilevel Tanpa Z | 46 |
| Tabel 4.9 | Uji Estimasi Parameter Analisis Regresi Multilevel Tanpa Z dan Tanpa X_2 | 47 |
| Tabel 4.10 | Uji Estimasi Parameter Analisis Regresi Multilevel Dengan Z..... | 47 |
| Tabel 4.11 | Nilai <i>Deviance</i> Setiap Model | 49 |
| Tabel 4.12 | Ketepatan Klasifikasi Untuk Analisis Regresi Logistik Biner..... | 50 |
| Tabel 4.13 | Ketepatan Klasifikasi Untuk Analisis Regresi Multilevel | 51 |

DAFTAR LAMPIRAN

| | |
|--------------------------------------------------------------------------------------------------------|----|
| Lampiran 1. Data Penelitian yang Digunakan pada Analisis Regresi Logistik Biner (Level 1) | 59 |
| Lampiran 2. Data Penelitian yang Digunakan pada Analisis Regresi Multilevel (Level 2)..... | 60 |
| Lampiran 3. Syntax <i>Software R</i> Untuk Analisis Regresi Logistik Biner..... | 61 |
| Lampiran 4. Syntax <i>Software R</i> Untuk Analisis Regresi 2-Level Tanpa Variabel Z..... | 61 |
| Lampiran 5. Syntax <i>Software R</i> Untuk Analisis Regresi 2-Level Dengan Variabel Z | 61 |
| Lampiran 6. Output <i>Software R</i> Untuk Analisis Regresi Logistik Biner..... | 62 |
| Lampiran 7. Output <i>Software R</i> Untuk Analisis Regresi 2-Level Tanpa Variabel Z..... | 63 |
| Lampiran 8. Output <i>Software R</i> Untuk Analisis Regresi 2-Level Dengan Variabel Z | 64 |

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Perkembangan ekonomi di Indonesia yang semakin pesat membuat kebutuhan rumah tangga semakin meningkat. Kurangnya pendapatan yang dihasilkan oleh suami sebagai kepala rumah tangga dan pencari nafkah membuat sebagian besar wanita ikut serta untuk bekerja guna memenuhi kebutuhan keluarga. Keadaan ekonomi keluarga yang semakin sulit sering kali memaksa beberapa anggota keluarga khususnya wanita untuk mencari nafkah, karena mengingat penghasilan suami yang tidak mencukupi untuk memenuhi kebutuhan hidup yang semakin sukar. Hal ini terlihat jelas pada keluarga dengan kondisi ekonomi menengah ke bawah, sehingga mendorong wanita untuk ikut berperan dalam meningkatkan pendapatan keluarganya dengan bekerja di sektor publik. Selain itu, sebagian wanita dari keluarga dengan kondisi ekonomi menengah ke atas yang juga ikut terjun ke dalam dunia kerja.

Pandangan yang memisahkan peran antara pria dan wanita sudah tidak relevan lagi, salah satunya ditunjukkan dengan semakin meningkatnya persentase keterlibatan wanita dalam dunia kerja (Alzeta dan Hidayati, 2009). Partisipasi wanita saat ini bukan sekedar menuntut persamaan hak tetapi juga menyatakan fungsinya yang mempunyai arti bagi pembangunan dalam masyarakat Indonesia. Melihat potensi wanita sebagai sumber daya manusia dapat berupaya untuk menyertakan wanita dalam proses pembangunan bukan hanya merupakan perikemanusiaan belaka, tetapi juga merupakan tindakan efisien karena dengan mengikutsertakan wanita dalam proses pembangunan berarti memberikan pengaruh yang positif terhadap lajunya pertumbuhan ekonomi (Sajogyo, 1983). Kecenderungan wanita untuk ikut bekerja menimbulkan banyak implikasi, antara lain tidak dapat menyeimbangkan kedua peran yaitu sebagai ibu rumah tangga dan sebagai wanita karir, sehingga menimbulkan berbagai konflik. Konflik yang terjadi adalah terjadinya benturan antara tanggungjawab wanita sebagai pekerja dan tanggungjawab wanita sebagai ibu rumah tangga serta istri (Nelson dan Quick, 2010).

Keputusan untuk mengambil dua peran yang berbeda yaitu sebagai ibu rumah tangga dan sebagai wanita karir tentu diikuti dengan tuntutan dari dalam diri sendiri dan pemenuhan kebutuhan kondisi ekonomi keluarga. Tuntutan dari diri sendiri dan pemenuhan kebutuhan hidup ini menyerukan hal yang sama yaitu keberhasilan dalam menjalankan dua peranan tersebut. Namun, bukanlah hal yang mudah bagi setiap wanita yang idealnya bisa menjalani semua peran dengan baik dan sempurna. Banyak dari wanita yang mempunyai peran ganda merasakan bahwa secara operasional kesulitan untuk membagi waktu antara urusan rumah tangga dan urusan pekerjaan. Akibat dari kesulitan yang sering dihadapi oleh wanita berperan ganda adalah keberhasilan untuk membagi setengah-setengah pada masing-masing peran atau hanya berhasil disalah satu peran saja dan peran yang lain dinomorduakan kemudian terbengkalai.

Sesuai data Survei Tenaga Kerja Nasional (Sakernas) tahun 2013, jumlah penduduk wanita yang bekerja sebagai buruh/ karyawan/ pegawai di Indonesia dari tahun 1980 mencapai 32,65 persen cenderung terus meningkat hingga 52,36 persen pada tahun 2013. Data tersebut juga didukung dengan meningkatnya jumlah penduduk di Indonesia yang semakin meningkat dari tahun ke tahun. Hal ini mengindikasikan bahwa kesempatan bekerja untuk wanita terus meningkat, sehingga kaum wanita semakin memiliki kesempatan untuk ikut andil dalam dunia pekerjaan. Melihat kecenderungan yang seperti ini, diprediksikan jumlah persentase wanita yang bekerja akan mengalami peningkatan pada masa-masa mendatang.

Jumlah wanita yang berada dalam dunia kerja (bekerja maupun sedang secara aktif mencari pekerjaan) telah meningkat secara drastis. Khususnya, wanita yang sudah menikah dan memiliki anak, memasuki dunia kerja dengan jumlah yang terus bertambah (Dubeck dan Borman, 1996). Beberapa penelitian sebelumnya tentang wanita bekerja, antara lain telah dilakukan oleh Anindita (2016) yaitu pemodelan faktor-faktor yang mempengaruhi ibu rumah tangga untuk bekerja di Provinsi Papua menggunakan analisis regresi logistik biner dan penelitian yang dilakukan oleh Majid (2012) yaitu faktor-faktor yang mempengaruhi keputusan wanita berstatus menikah untuk bekerja (studi kasus : kota Semarang) dengan menggunakan regresi logistik biner.

Analisis regresi logistik dibagi menjadi beberapa jenis data pada responnya, diantaranya yaitu respon biner, multinomial, ordinal dan poisson. Analisis regresi logistik yang paling sederhana digunakan adalah analisis regresi logistik biner. Regresi logistik biner digunakan untuk menganalisa hubungan antara satu variabel respon dan beberapa variabel independen, dengan variabel responnya berupa data kualitatif dikotomi yaitu bernilai 1 untuk menyatakan keberadaan sebuah karakteristik dan bernilai 0 untuk menyatakan ketidakberadaan sebuah karakteristik (Agresti, 2007). Kini, data dengan respon biner tidak hanya digunakan pada regresi logistik saja, namun sudah mulai berkembang pada metode regresi lainnya.

Berbagai kasus, seperti kasus sosial, pendidikan dan kesehatan, sering kali dijumpai data populasi yang memiliki struktur hirarki. Struktur data hirarki biasanya berasal dari populasi yang bertingkat atau berjenjang (*hierarchy*) dan berkelompok (*cluster*). Data yang memiliki struktur hirarki adalah data yang terjadi karena individu-individu terkumpul dalam kelompok-kelompok sosialnya. Data yang memiliki struktur hirarki, sebenarnya terdapat efek dari kelompok yang membawahi unit-unit yang diteliti (Hox, 2002). Sebagai contoh struktur data hirarki pada kasus sosial adalah penduduk yang terkelompok dalam suatu wilayah, pada kasus pendidikan adalah siswa yang terkelompok dalam suatu sekolah dan pada kasus kesehatan adalah pasien yang terkelompok dalam suatu rumah sakit.

Salah satu data yang memiliki struktur data berjenjang adalah data survei yang dilakukan oleh Indonesian Family Life Survey (IFLS) dari tahun 1993 meliputi survei tentang lingkungan sosial, ekonomi, kesehatan, dan sebagainya. IFLS melakukan survei pada rumah tangga dan survei pada komunitas/fasilitas. Untuk survei rumah tangga, yang diwawancarai adalah individu-individu dalam suatu rumah tangga, sedangkan untuk survei komunitas/fasilitas berdasarkan pada Kecamatan. Dengan kata lain, dapat dikatakan sebagai individu yang berkelompok dalam Kecamatan. Hal ini mengindikasikan bahwa terdapat hubungan antara individu dan lingkungannya, misalnya Kecamatan, dalam arti bahwa suatu individu dipengaruhi oleh Kecamatan yang merupakan tempat dimana mereka tinggal, dan sifat atau kebiasaan dalam Kecamatan terbentuk dari individu-individu yang berdomisili dalam Kecamatan tersebut.

Suatu metode yang digunakan untuk menyelesaikan kasus berstruktur hirarki adalah model multilevel. Model multilevel merupakan teknik statistik yang telah mengalami pengembangan dari regresi klasik/ sederhana. Pengembangan tersebut didasari karena dalam penelitian diberbagai kasus seperti kasus sosial sering dijumpai perbedaan latar belakang pada responden yang diteliti, sehingga data tersebut memiliki struktur yang bertingkat atau berjenjang (*hierarchy*) dan berkelompok (*cluster*). Perbedaan tersebut muncul karena data yang diperoleh dari survei yang dilakukan menggunakan penarikan random sampling bertahap (*multistage random sampling*) dan adanya hubungan antara variabel pada tingkat yang berbeda (Goldstein, 1995) sehingga kondisi tersebut akan menghasilkan data yang berstruktur hirarki (Hox, 2002). Goldstein (1995) memperkenalkan pengembangan dari regresi biasa untuk mengatasi permasalahan yang disebabkan dari data yang berstruktur hirarki yaitu analisis *Multilevel Modeling*.

Tingkatan struktur hirarki dalam model multilevel didefinisikan sebagai level. Model multilevel yang paling sederhana adalah model dengan 2-level, yaitu hanya terdapat 2-level pada struktur data hirarki. Tingkat yang paling rendah yaitu individu disebut level 1 dan tingkat yang lebih tinggi yaitu Kecamatan disebut level 2. Ada banyak penelitian-penelitian sebelumnya tentang model multilevel, diantaranya telah dilakukan oleh Poedjiati (2009) yaitu perbandingan regresi logistik dan model multilevel pada data ordinal dan biner untuk mengetahui variabel yang mempengaruhi hasil belajar mengajar kalkulus I, yang berisi tentang pengamatan pada data pendidikan menggunakan model 2 level. Penelitian lainnya telah dilakukan oleh Antika (2011) yaitu analisis regresi multilevel terhadap faktor-faktor yang mempengaruhi nilai UASBN SD/MI, yang berisi tentang pengamatan pada data pendidikan menggunakan model 2 level.

Model multilevel merupakan bagian dari *Generalized Linear Mixed Model* (GLMM), yaitu variabel respon dipengaruhi oleh faktor tetap dan faktor acak. Penelitian tentang GLMM telah banyak dilakukan, diantaranya oleh Fong, dkk (2009) yaitu tentang *bayesian inference for generalized linear mixed models* yang diterapkan pada data longitudinal dengan pendekatan laplace bersarang terintegrasi dan dilakukan oleh Khan dan Shaw (2011) yaitu *multilevel logistic regression analysis applied to binary contraceptive prevalence data*, yang berisi tentang

pengamatan pada survei kesehatan di Bangladesh menggunakan model 3 level menggunakan pendekatan *Penalized Quasi-Likelihood*. Salah satu penelitian yang digunakan sebagai pedoman dalam penelitian ini adalah penelitian yang telah dilakukan oleh Guo dan Zhao (2000) yaitu *multilevel modeling for binary data*, yang berisi tentang pengamatan pada data pendidikan di United States menggunakan model 2 level. Oleh karena itu, dalam penelitian ini akan dilakukan analisis tentang status bekerja ibu rumah tangga dengan menggunakan model 2 level.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diperoleh, rumusan masalah yang akan dibahas dalam penelitian ini adalah mengestimasi parameter untuk model 2 level dengan respon biner dan pemodelan wanita berstatus menikah yang memutuskan untuk bekerja menggunakan model 2 level dengan respon biner.

1.3 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah untuk memecahkan masalah yang telah diidentifikasi sebelumnya, adapun tujuan penelitian ini sebagai berikut.

1. Mendapatkan estimasi parameter untuk model 2 level dengan respon biner.
2. Menentukan model 2 level dengan respon biner yang sesuai untuk status bekerja ibu rumah tangga.

1.4 Manfaat Penelitian

Manfaat yang dapat diperoleh dari penelitian ini adalah menambah pengetahuan tentang pemodelan yang sesuai untuk kasus status bekerja ibu rumah tangga dengan menerapkan estimasi parameter dalam model regresi 2 level dengan respon biner. Selain itu, manfaat untuk pemerintah Indonesia adalah dapat memberikan kontribusi untuk lebih memperhatikan kualitas tenaga kerja wanita dengan mempertimbangkan usia, pendidikan terakhir dan pengeluaran rumah tangga, serta rata-rata gaji wanita.

1.5 Batasan Masalah

Batasan masalah pada penelitian tentang status bekerja ibu rumah tangga adalah hanya mengambil 5 variabel yaitu variabel status bekerja ibu rumah tangga, variabel usia, variabel pendidikan terakhir dan variabel pengeluaran rumah tangga pada level 1 (individu), serta variabel rata-rata gaji wanita pada level 2 (Kecamatan). Variabel-variabel tersebut dipilih karena data pada variabel tersebut lengkap dan dianggap memiliki hubungan dengan status bekerja ibu rumah tangga. Selain itu, batasan masalah dalam penelitian ini adalah hanya untuk wanita berstatus menikah atau pernah menikah yang ada di Indonesia. Hal ini diteliti karena semakin meningkatnya persentase wanita bekerja dari tahun ke tahun, terutama wanita berstatus menikah atau pernah menikah.

BAB 2

TINJAUAN PUSTAKA

2.1 Multikolinearitas

Adanya korelasi yang tinggi antar variabel prediktor dinamakan multikolinieritas. Uji multikolinearitas bertujuan untuk menguji adanya korelasi antar variabel independen (variabel bebas) dalam model regresi (Gujarati, 2004). Untuk melihat adanya multikolinieritas dapat digunakan VIF (*Variance Inflation Factor*) dengan rumus sebagai berikut :

$$VIF = \frac{1}{1 - R_i^2} \quad (2.1)$$

dimana :

$VIF < 10$: Mengindikasikan tidak ada korelasi yang signifikan antar variabel independen.

$VIF \geq 10$: Mengindikasikan bahwa ada salah satu variabel independen merupakan fungsi dari variabel prediktor yang lain.

2.2 Model Regresi

Model regresi merupakan model yang digunakan untuk melihat hubungan antara satu variabel tak bebas (variabel dependen), atau disebut juga dengan variabel respon, dengan beberapa variabel bebas (variabel independen), atau disebut juga dengan variabel penjelas (Draper and Smith, 1992). Model regresi linear adalah model yang menunjukkan hubungan yang linear antara variabel respon dan variabel penjelas.

Terdapat dua persamaan model dalam model regresi linear, yaitu regresi linier sederhana (*simple regression*) dan regresi linear berganda (*multiple regression*). Jika persamaan model regresi hanya terdiri dari satu variabel bebas, maka model tersebut disebut dengan regresi linear sederhana. Sedangkan, jika persamaan model regresi terdapat lebih dari satu variabel bebas, maka model tersebut disebut dengan regresi linear berganda.

Secara umum, model regresi linear sebagai berikut.

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_i X_i + \varepsilon_i \quad (2.2)$$

Dimana :

i : Banyaknya variabel independen, $i = 1, 2, \dots, p$

Y : Variabel respon

X_i : Variabel penjelas

ε_i : Error

β_i : Parameter atau disebut juga dengan koefisien regresi.

Nilai koefisien regresi pada umumnya tidak diketahui, sehingga harus dilakukan pendugaan. Untuk menduga koefisien regresi pada persamaan (2.2) dapat digunakan metode kuadrat terkecil (OLS) atau metode kemungkinan maksimum (MLE) (Kutner et. al, 2004).

Untuk data berstruktur hirarki, misalnya struktur 2 level, persamaan (2.2) tidak mencakup informasi pada level 2 atau informasi diabaikan. Sehingga, untuk mencakup informasi pada level 2, semua variabel termasuk error yang terdapat pada level 2 harus dimasukkan ke dalam persamaan regresinya.

2.3 Model Regresi Multilevel

Model regresi multilevel merupakan bagian dari model linear campuran (*Mixed effect model* atau *Generalized Linear Mixed Model (GLMM)*), yaitu suatu model yang menggabungkan komponen tetap atau efek tetap (*fixed effect*) dan komponen acak atau efek acak (*random effect*) ke dalam satu persamaan model. Model yang termasuk ke dalam model linear campuran adalah model-model data hasil pengukuran berulang (*repeated measures*), model data longitudinal dan model multilevel. Model multilevel mulai mendapat perhatian setelah dikembangkan oleh Goldstein (1995), akan tetapi model ini jarang digunakan karena melibatkan perhitungan yang cukup rumit. Estimasi parameter untuk model multilevel diantaranya dengan menggunakan *Maximum Likelihood (ML)* atau *Restricted Maximum Likelihood (REML)*.

Model multilevel merupakan sebuah model yang digunakan pada data berjenjang (*hierarchy*). Data berjenjang seringkali ditemukan pada penelitian-penelitian survei dimana unit-unit analisisnya berasal dari kelompok-kelompok (*cluster*), atau data yang diambil melalui penarikan sampel bertahap (*cluster sampling*). Misalnya, dalam pengambilan sampel menggunakan metode sampling

satu tahap (*Single Stage Cluster Sampling*), dimana unit-unit sampling yang berasal dari kelompok diperhitungkan keberadaanya dalam analisis, sehingga dalam hal ini model yang sesuai adalah model multilevel. Unit-unit sampling yang ada dalam kelompok disebut level rendah dan kelompok-kelompok disebut level tinggi. Banyaknya unit-unit analisis dalam kelompok bisa sama atau berbeda untuk setiap kelompok.

Secara umum, model regresi multilevel mempunyai struktur data berjenjang (hierarki) sebagai berikut.

1. Sebuah variabel tak bebas (variabel dependen) yang diukur pada level rendah.
2. Beberapa variabel bebas (variabel independen) yang diukur pada setiap level.

2.3.1 Model Regresi 2 Level

Suatu model regresi multilevel yang sederhana hanya terdiri dari 2 level. Misalkan, diberikan data dalam J kelompok dan jumlah yang berbeda dari individu n_j dalam setiap kelompok. Dalam individu (level 1), terdapat variabel tak bebas/respon Y_{ij} dan variabel bebas/penjelas X_{ij} , serta pada level kelompok terdapat variabel bebas Z_j . Sehingga, terdapat persamaan model regresi terpisah pada setiap kelompok. Model regresi 2 level dengan satu variabel bebas sebagai berikut.

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \varepsilon_i \quad (2.3)$$

Dimana :

indeks i menyatakan individu pada tingkat Level 2 ke- j ($i = 1, 2, \dots, n_j$)

indeks j menyatakan tingkat Level 2 ($j = 1, 2, \dots, m$)

Dalam persamaan (2.3), β_0 adalah intersep dalam regresi klasik, β_0 adalah *slope* regresi untuk variabel penjelas X , dan ε_i merupakan residual error. Koefisien regresi β dengan j untuk kelompok, dimana mengindikasikan bahwa koefisien regresi mungkin bervariasi setiap kelompok. Variasi dalam koefisien regresi β_j dimodelkan dengan variabel penjelas dan residual acak pada level kelompok, sehingga diperoleh persamaan sebagai berikut.

$$\beta_{0j} = \beta_{00} + \beta_{01}Z_j + u_{0j} \quad (2.4)$$

$$\beta_{1j} = \beta_{10} + \beta_{11}Z_j + u_{1j} \quad (2.5)$$

Persamaan (2.3) disebut sebagai model pada level 1 dan persamaan (2.4) dan persamaan (2.5) disebut sebagai model pada level 2. Selanjutnya, persamaan (2.4)

dan persamaan (2.5) disubstitusikan ke dalam persamaan (2.3), sehingga diperoleh model regresi multilevel sebagai berikut.

$$\begin{aligned} Y_{ij} &= \beta_{00} + \beta_{01}Z_j + u_{0j} + (\beta_{10} + \beta_{11}Z_j + u_{1j})X_{ij} + \varepsilon_i \\ Y_{ij} &= \beta_{00} + \beta_{01}Z_j + u_{0j} + \beta_{10}X_{ij} + \beta_{11}Z_jX_{ij} + u_{1j}X_{ij} + \varepsilon_i \\ Y_{ij} &= \beta_{00} + \beta_{01}Z_j + \beta_{10}X_{ij} + \beta_{11}Z_jX_{ij} + u_{0j} + u_{1j}X_{ij} + \varepsilon_i \end{aligned} \quad (2.6)$$

Dari persamaan (2.6), $\beta_{00} + \beta_{01}Z_j + \beta_{10}X_{ij} + \beta_{11}Z_jX_{ij}$ merupakan bagian dari efek tetap, sedangkan $u_{0j} + u_{1j}X_{ij} + \varepsilon_i$ merupakan bagian dari efek acak. Dari persamaan tersebut juga dapat terlihat bahwa nilai Y_{ij} secara umum dapat diprediksi oleh Z_j dan juga dapat diketahui bahwa hubungan fungsional antara Y_{ij} dan X_{ij} bergantung pada nilai Z_j . Sehingga persamaan (2.6) dapat dikatakan sebagai bagian dari model linear campuran.

Persamaan (2.6) dapat dituliskan dalam bentuk matriks dengan menotasikan variabel bebas pada efek tetap sebagai X dan variabel bebas pada efek acak sebagai Z , sehingga diperoleh :

$$Y_{ij} = X_{ij}\beta + Z_ju_j + \varepsilon_i \quad (2.7)$$

$i = 1, 2, \dots, n ; j = 1, 2, \dots, m$

Dimana :

$$\begin{aligned} \begin{pmatrix} Y_{ij} \\ (n_{ij} \times 1) \end{pmatrix} &= \begin{bmatrix} Y_{1j} \\ Y_{2j} \\ \vdots \\ Y_{n_{ij}j} \end{bmatrix}, \quad \begin{pmatrix} X_{ij} \\ (n \times m) \end{pmatrix} = \begin{bmatrix} 1 & X_{11} & \cdots & X_{n1} \\ 1 & X_{12} & \cdots & X_{n2} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & X_{1j} & \cdots & X_{nm} \end{bmatrix}, \quad \begin{pmatrix} Z_{ij} \\ (m \times 2) \end{pmatrix} = \begin{bmatrix} 1 & Z_1 \\ 1 & Z_2 \\ \vdots & \vdots \\ 1 & Z_m \end{bmatrix}, \\ \begin{pmatrix} \beta \\ (4 \times 1) \end{pmatrix} &= \begin{bmatrix} \beta_{00} \\ \beta_{01} \\ \beta_{10} \\ \beta_{11} \end{bmatrix}, \quad \begin{pmatrix} u_j \\ (2 \times 1) \end{pmatrix} = \begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix}, \quad \begin{pmatrix} \varepsilon_i \\ (n \times 1) \end{pmatrix} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \end{aligned}$$

Secara umum, ada lebih dari satu variabel penjelas pada level 1 atau level terendah dan juga lebih dari satu variabel penjelas pada level 2 atau lebih tinggi. Diberikan bahwa P variabel penjelas X pada level 1 atau level terendah, diindikasikan dengan p ($p = 1, 2, \dots, P$), dan Q variabel penjelas Z pada level 2 atau lebih tinggi, diindikasikan dengan q ($q = 1, 2, \dots, Q$). Sehingga, diperoleh persamaan yang lebih umum sebagai berikut.

Model pada level 1

$$Y_{ij} = \beta_{0j} + \sum_{p=1}^P \beta_{pij} X_{pij} + \varepsilon_i \quad (2.8)$$

Model pada level 2

$$\beta_{0j} = \beta_{00} + \sum_{q=1}^Q \beta_{0q} Z_{qj} + u_{0j} \quad (2.9)$$

$$\beta_{pij} = \beta_{q0} + \sum_{q=1}^Q \beta_{qj} Z_{qj} + u_{pj} \quad (2.10)$$

$i = 1, 2, \dots, n$; $j = 1, 2, \dots, m$; $p = 1, 2, \dots, P$; $q = 1, 2, \dots, Q$

Dimana :

Y_{ij} : Variabel dependen untuk unit ke-i level 1 dalam unit ke-j level 2

β_{0j} : *Random intercept* untuk unit ke-j level 2

β_{0q} : Efek tetap (*fixed effects*) untuk variabel independen ke-q level 2

β_{pij} : Efek tetap (*fixed effects*) untuk variabel independen ke-p pada unit ke-i level 1 dalam unit ke-j level 2

β_{q0} : Efek tetap (*fixed effects*) untuk variabel independen ke-q level 2

β_{qj} : Efek tetap (*fixed effects*) untuk variabel independen ke-q pada unit ke-j level 2

X_{pij} : Variabel independen ke-p pada unit ke-i level 1 dalam unit ke-j level 2

Z_{qj} : Variabel independen ke-q untuk unit ke-j level 2

ε_{ij} : Residual error untuk unit ke-i level 1 dalam unit ke-j level 2 (residual level 1), diasumsikan berdistribusi $N(0, \sigma_\varepsilon^2)$

u_{0j} : Residual error untuk unit ke-j level 2 (residual level 1), diasumsikan berdistribusi $N(0, \sigma_u^2)$

u_{pj} : Residual error untuk unit ke-i level 1 dalam unit ke-j level 2 (residual level 2), diasumsikan berdistribusi $N(0, \sigma_u^2)$

Persamaan (2.9) dan persamaan (2.10) disubstitusikan ke dalam persamaan (2.8), sehingga diperoleh model umum untuk regresi 2 level sebagai berikut.

$$Y_{ij} = \beta_{00} + \sum_{q=1}^Q \beta_{0q} Z_{qj} + \sum_{p=1}^P \beta_{q0} X_{pij} + \sum_{q=1}^Q \sum_{p=1}^P \beta_{pq} Z_{qj} X_{pij} + \sum_{p=1}^P u_{pj} X_{pij} + u_{0j} + \varepsilon_i \quad (2.11)$$

Dalam persamaan (2.11), β merupakan koefisien regresi, u merupakan istilah residual pada level kelompok, dan ε merupakan residual pada level individu.

Koefisien regresi diidentifikasi sebagai bagian tetap dari model karena bagian ini tidak dapat berubah, baik dari kelompok atau individu. Residual error diidentifikasi sebagai bagian acak dari model (Hox, 2002).

Struktur data untuk model regresi 2 level dapat dilihat pada Tabel 2.1 sebagai berikut.

Tabel 2.1 Struktur Data Model 2 Level

| Kelompok | Pengamatan | Variabel | Variabel | | | Variabel | | |
|----------|----------------|------------------|--------------------|-----|-------------------|--------------------|-----|-----------------|
| | | Dependen | Independen Level 1 | | | Independen Level 2 | | |
| | | Y | X ₁ | ... | X _p | Z ₁ | ... | Z _q |
| 1 | 1 | Y ₁₁ | X ₁₁₁ | | X _{p11} | | | |
| | 2 | Y ₂₁ | X ₁₂₁ | | X _{p21} | Z ₁₁ | ... | Z _{q1} |
| | ⋮ | ⋮ | ⋮ | ... | ⋮ | | | |
| | n ₁ | Y _{n11} | X _{1n11} | | X _{pn11} | | | |
| 2 | 1 | Y ₁₂ | X ₁₁₂ | | X _{p12} | | | |
| | 2 | Y ₂₂ | X ₁₂₂ | | X _{p22} | Z ₁₂ | ... | Z _{q2} |
| | ⋮ | ⋮ | ⋮ | ... | ⋮ | | | |
| | n ₂ | Y _{n22} | X _{1n22} | | X _{pn22} | | | |
| ⋮ | ⋮ | ⋮ | ⋮ | ... | ⋮ | ⋮ | ... | ⋮ |
| m | 1 | Y _{1m} | X _{11m} | | X _{p1m} | | | |
| | 2 | Y ₂₁ | X _{12m} | | X _{p2m} | Z _{1m} | ... | Z _{qm} |
| | ⋮ | ⋮ | ⋮ | ... | ⋮ | | | |
| | n _m | Y _{nmm} | X _{1nmm} | | X _{pnmm} | | | |

2.3.2 Model Regresi 2 Level Dengan *Random Intercept*

Salah satu jenis model regresi multilevel adalah model regresi multilevel dengan *random intercept*. Dalam model ini, diasumsikan hanya koefisien intersep yang bersifat acak (random). Model regresi multilevel dengan *random intercept* ini memiliki koefisien regresi yang berbeda untuk setiap unit level dua.

Misalkan terdapat P variabel bebas pada level 1 dan terdapat Q variabel bebas pada level 2, sehingga model regresi multilevel menggunakan *random intercept model* dapat dilihat sebagai berikut (Raudenbush dkk, 2001).

Model pada level 1 :

$$Y_{ij} = \beta_{0j} + \sum_{p=1}^P \beta_{pi} X_{pi} + \varepsilon_i \quad (2.12)$$

Model pada level 2 :

$$\beta_{0j} = \beta_{00} + \sum_{q=1}^Q \beta_{0q} Z_{qj} + u_{0j} \quad (2.13)$$

$i = 1, 2, \dots, n ; j = 1, 2, \dots, m ; p = 1, 2, \dots, P ; q = 1, 2, \dots, Q$

Dimana :

Y_{ij} : Variabel dependen untuk unit ke-i level 1 dalam unit ke-j level 2

β_{0j} : *Random intercept* untuk unit ke-j level 2

β_{pi} : Efek tetap (*fixed effects*) untuk variabel independen ke-p pada unit ke-i level 1 dalam unit ke-j level 2

X_{pi} : Variabel independen ke-p pada unit ke-i level 1 dalam unit ke-j level 2

β_{00} : Intersep tetap (*fixed intercept*)

β_{0q} : Efek tetap (*fixed effects*) untuk variabel independen ke-q level 2

Z_{qj} : Variabel independen ke-q untuk unit ke-j level 2

ε_i : Residual untuk unit ke-i level 1 dalam unit ke-j level 2 (residual level 1), diasumsikan berdistribusi $N(0, \sigma_\varepsilon^2)$

u_{0j} : Residual error untuk unit ke-j level 2 (residual level 1), diasumsikan berdistribusi $N(0, \sigma_u^2)$

Dalam model pada persamaan (2.12) dan persamaan (2.13), notasi $i = 1, 2, \dots, n_j$ menyatakan unit-unit pada level 1 yang tersarang dalam unit ke-j pada level 2 dan $j = 1, 2, \dots, m$ menyatakan unit-unit pada level 2. Total observasi pada level 1 dalam seluruh unit pada level 2 sebagai berikut.

$$n = \sum_{j=1}^m n_j \quad (2.14)$$

Model pada persamaan (2.14) dapat disubstitusikan ke dalam model pada persamaan (2.13), sehingga model regresi 2-level dengan *random intercept* sebagai berikut.

$$Y_{ij} = \beta_{00} + \sum_{q=1}^Q \beta_{0q} Z_{qj} + \sum_{p=1}^P \beta_{pi} X_{pi} + u_{0j} + \varepsilon_i \quad (2.15)$$

Model pada persamaan (2.15) merupakan model kombinasi.

Umumnya, asumsi yang digunakan model regresi multilevel adalah bahwa residual pada level 1 atau ε_i level terendah mengikuti distribusi normal dengan rata-rata 0 dan varians σ^2 dalam semua kelompok. Residual level kedua u_{0j} diasumsikan independen dari ε_i error level terendah dan mengikuti distribusi normal dengan rata-rata 0. Kebanyakan *software* multilevel mengasumsikan bahwa varians dari residual error ε_i adalah sama dalam semua kelompok (Hox, 2002).

2.4 Model Regresi Logistik Biner

Regresi logistik digunakan untuk mencari hubungan variabel dependen (Y) yang bersifat *dichotomous* (berskala nominal atau ordinal dengan dua kategori) atau *polychotomous* (mempunyai skala nominal atau ordinal dengan lebih dari dua kategori) dengan satu atau lebih variabel independen (X) yang bersifat kontinu atau kategorik (Agresti, 2007). Salah satu regresi logistik yang paling sederhana digunakan adalah regresi logistik biner. Regresi logistik biner merupakan suatu metode analisis data yang digunakan untuk mencari hubungan antara variabel dependen (Y) yang bersifat *biner* atau dikotomis dengan variabel independen (X) yang bersifat polikotomis (Hosmer dan Lemeshow, 1989).

Data variabel dependen yang digunakan dalam regresi logistik biner adalah data dengan skala nominal dengan hanya berupa 2 kategori yaitu “sukses” atau “gagal” misalnya: ya-tidak, benar-salah, hidup-mati, hadir-absen, laki-wanita, dan seterusnya. Sedangkan data variabel independen dapat berupa data dengan skala ordinal (seringkali digunakan pada kasus-kasus/penelitian sosial kemasyarakatan) ataupun data dengan skala rasio (seringkali dijumpai pada penelitian industri). *Outcome* dari variabel dependen Y terdiri dari 2 kategori yaitu “sukses” dan “gagal” yang dinotasikan dengan $Y=1$ (sukses) dan $Y=0$ (gagal). Dalam keadaan demikian, variabel Y mengikuti distribusi Bernoulli untuk setiap observasi tunggal. Fungsi probabilitas untuk setiap observasi diberikan sebagai berikut.

$$f(Y) = p^Y (1 - p)^{1-Y} \quad (2.16)$$

Persamaan (2.15) dapat menunjukkan bahwa regresi logistik biner merupakan keluarga eksponensial dengan syarat

$$Y(Y; \theta) = a(\theta)b(Y) \exp(\theta)d(Y) \quad (2.17)$$

Jika X_1, X_2, \dots, X_k merupakan variabel random yang diduga mempengaruhi besar kecilnya π , dimana k adalah banyaknya variabel independen. Sehingga, fungsi linknya sebagai berikut.

$$c(\pi) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (2.18)$$

dan model regresi logistik biner diperoleh sebagai berikut.

$$\ln\left(\frac{\pi}{1-\pi}\right) = X^T \beta \text{ dapat ditulis juga } \ln\left(\frac{\pi(X)}{1-\pi(X)}\right) = X^T \beta$$

Variabel dependen dalam regresi logistik dimodelkan sebagai $Y = \pi(X) + \varepsilon$ dimana ε mempunyai salah satu dari kemungkinan dua nilai yaitu $\varepsilon = 1 - \pi(X)$ dengan peluang $\pi(X)$ jika $Y=1$ dan $\varepsilon = -\pi(X)$ dengan peluang $1-\pi(X)$ jika $Y=0$ dan mengikuti distribusi binomial $(1, \pi(X))$ dengan rata-rata nol dan varians $(\pi(X))(1-\pi(X))$.

Struktur data untuk model regresi logistik biner dapat dilihat pada Tabel 2.2 sebagai berikut.

Tabel 2.2 Struktur Data Regresi Logistik Biner

| Pengamatan ke-i | X_1 | X_2 | ... | X_k | Y |
|-----------------|----------|----------|-----|----------|----------|
| 1 | X_{11} | X_{21} | ... | X_{k1} | Y_1 |
| 2 | X_{12} | X_{22} | ... | X_{k2} | Y_2 |
| \vdots | \vdots | \vdots | ... | \vdots | \vdots |
| n | X_{1n} | X_{2n} | ... | X_{kn} | Y_n |

2.5 Model Regresi Multilevel Respon Biner

Model regresi multilevel dengan variabel respon (dependen) berupa data biner atau hanya terdiri dari dua kategori, maka estimasi parameter bisa dilakukan menggunakan estimasi maksimum likelihood dengan pendekatan suatu metode tertentu (Goldstein, 1999).

Dalam model regresi, apabila variabel respon berupa biner atau *dichotomous* biasanya digunakan model regresi logistik yang dalam estimasi parameternya harus menggunakan suatu fungsi penghubung (*link function*). Hal tersebut juga sama diterapkan dalam model multilevel. Apabila variabel respon berdistribusi binomial dengan parameter proporsi (π_{ij}) , maka fungsi penghubung yang digunakan adalah logit ($\log\{\pi/(1-\pi)\}$) sehingga modelnya disebut dengan model logistik (Hox, 2002).

Secara umum, model 2 level dengan respon biner dapat dituliskan sebagai berikut (Guo dan Zhao, 2000).

$$\log \left[\frac{\pi_{ij}}{(1 - \pi_{ij})} \right] = \beta_0 + \beta_1 X_{li} + u_j \quad (2.19)$$

Dimana u_j merupakan efek acak pada level 2, tanpa u_j , persamaan (2.19) menjadi model regresi logistik standar. Bergantung pada u_j , Y_{ij} diasumsikan independen. Seperti dalam kasus pada model linear multilevel, u_j diasumsikan berdistribusi normal, dengan nilai yang diharapkan 0 dan varians σ_u^2 . Model (2.19) seringkali

dideskripsikan sebagai alternatif dalam literatur pada model multilevel dari persamaan berikut.

$$\log \left[\frac{\pi_{ij}}{(1 - \pi_{ij})} \right] = \beta_{0j} + \beta_1 X_{1i} + \varepsilon_i \text{ (model level 1)} \quad (2.20)$$

dan

$$\beta_{0j} = \beta_0 + u_j \text{ (model level 2)} \quad (2.21)$$

Secara umum rumusan matematis untuk model *random-intercept* dua level dengan respon biner sebagai berikut

$$\log \left[\frac{\pi_{ij}}{(1 - \pi_{ij})} \right] = \beta_0 + \beta_1 X_{1i} + u_j + \varepsilon_i \quad (2.22)$$

Dimana u_j merupakan error pada level 2 yang berdistribusi normal dengan rata-rata 0 dan varians σ_u^2 dan independen terhadap ε_i .

2.6 Estimasi Parameter Model Regresi Multilevel Dengan Respon Biner

Estimasi parameter dalam model regresi multilevel dengan respon biner menggunakan *Maximum Likelihood Estimation* (MLE) dengan pendekatan integral kuadrat Gauss-Hermite Quadrature.

2.6.1 Generalized Linear Mixed Model (GLMM)

Metode yang digunakan untuk menentukan kesesuaian model dalam model regresi multilevel, yaitu *Generalized Linear Mixed Model* (GLMM) (Agresti, 2007). GLMM merupakan model dimana variabel respon dipengaruhi oleh faktor tetap dan faktor acak. Parameter dari faktor tetap atau faktor acak pada model tersebut tidak diketahui nilainya, sehingga harus dilakukan pendugaan. Bentuk umum dari GLMM sebagai berikut.

$$Y = X\beta + u + \varepsilon \quad (2.23)$$

2.6.2 Maximum Likelihood Estimation (MLE)

Vektor \mathbf{Y} dalam model multilevel linier berdistribusi normal yang merupakan kombinasi linier dari koefisien acak. Model multilevel non linier atau lebih khususnya adalah model dengan respon biner merupakan model regresi yang tidak dapat dijelaskan sebagai kombinasi linier dari koefisiennya sehingga dengan demikian vektor \mathbf{Y} tidak lagi berdistribusi normal. Estimasi parameter dalam regresi

logistik dilakukan dengan metode *Maximum Likelihood Estimation* (MLE). Metode tersebut mengestimasi parameter β dengan cara memaksimumkan fungsi likelihood dan mensyaratkan bahwa data harus mengikuti suatu distribusi tertentu. Setiap pengamatan dalam regresi logistik mengikuti distribusi bernoulli sehingga dapat ditentukan fungsi likelihoodnya.

Jika X_i dan Y_i adalah pasangan variabel bebas dan terikat pada pengamatan ke- i dan diasumsikan bahwa setiap pasangan pengamatan saling independen dengan pasangan pengamatan lainnya, $i = 1, 2, \dots, n$ maka fungsi probabilitas untuk setiap pasangan sebagai berikut

$$f(Y_i) = \pi(X_i)^{Y_i} (1 - \pi(X_i))^{1-Y_i}; \quad Y_i = 0, 1 \quad (2.24)$$

Setiap pasangan pengamatan diasumsikan independen sehingga fungsi likelihoodnya merupakan gabungan dari fungsi distribusi masing-masing pasangan sebagai berikut.

$$L = \prod_{i=1}^n f(Y_i) = \prod_{i=1}^n \pi(X_i)^{Y_i} (1 - \pi(X_i))^{1-Y_i} \quad (2.25)$$

Secara umum, persamaan (2.25) tidak bisa diselesaikan kecuali melalui prosedur iterasi dengan menggunakan integrasi numerik. Dalam model multilevel prosedur integrasi numerik yang digunakan adalah prosedur Gauss-Hermite Quadrature untuk menghitung integral secara numerik dengan persamaan sebagai berikut.

$$\int_{-\infty}^{\infty} (g(v)) \exp(-v^2) dv \approx \sum_{k=1}^m w_k g(v_k) \quad (2.26)$$

dimana :

w_k : Bobot quadrature

v_k : Titik quadrature

Fungsi likelihood tersebut lebih mudah dimaksimumkan dalam bentuk $\ln L$ dengan persamaan sebagai berikut.

$$L = \ln L = \ln \left[\prod_{i=1}^n \left(\frac{\pi(X_i)}{1 - \pi(X_i)} \right)^{Y_i} (1 - \pi(X_i)) \right] \quad (2.27)$$

Nilai β maksimum didapatkan melalui turunan L terhadap β dan hasilnya adalah sama dengan nol.

- Turunan Pertama

$$\frac{\partial L}{\partial \beta} = 0 = g(\beta) \quad (2.28)$$

merupakan vektor gradien dari L , sehingga

$$\mathbf{g}(\boldsymbol{\beta}) = \frac{\partial L}{\partial \boldsymbol{\beta}^T} = \mathbf{0} \quad (2.29)$$

Estimasi varians dan kovarians dikembangkan melalui teori MLE (*Maximum Likelihood Estimation*) dari koefisien parameternya (Hosmer dan Lemeshow, 1989). Teori tersebut menyatakan bahwa estimasi varians kovarians didapatkan melalui turunan kedua L , yaitu matriks Hessian dari $\boldsymbol{\beta}$.

- Turunan Kedua

$$\frac{\partial L}{\partial \boldsymbol{\beta}^T \boldsymbol{\beta}} = \mathbf{H}(\boldsymbol{\beta}) \quad (2.30)$$

$$\mathbf{H}(\boldsymbol{\beta}) = \begin{bmatrix} \frac{\partial^2 \ln L}{\partial \beta_0^2} & \frac{\partial^2 \ln L}{\partial \beta_0 \partial \beta_1} & \frac{\partial^2 \ln L}{\partial \beta_0 \partial \beta_2} & \dots & \frac{\partial^2 \ln L}{\partial \beta_0 \partial \beta_k} \\ \frac{\partial^2 \ln L}{\partial \beta_1^2} & \frac{\partial^2 \ln L}{\partial \beta_1 \partial \beta_2} & \dots & \frac{\partial^2 \ln L}{\partial \beta_1 \partial \beta_k} \\ \frac{\partial^2 \ln L}{\partial \beta_2^2} & \dots & \frac{\partial^2 \ln L}{\partial \beta_2 \partial \beta_k} \\ \vdots & \ddots & \vdots \\ \text{simetris} & & \frac{\partial^2 \ln L}{\partial \beta_k^2} \end{bmatrix}$$

Hasil yang diperoleh tidak *close form* karena nilai $\boldsymbol{\beta}$ berbentuk skalar, sehingga untuk mencari nilai estimasi $\boldsymbol{\beta}$ digunakan iterasi numerik.

Metode *Maximum Likelihood Estimation* memerlukan nilai awal yang baik untuk parameter-parameternya. Parameter untuk model non linier diduga dengan menggunakan *Generalized Least Square* (GLS) yang kemudian dijadikan sebagai nilai awal dari koefisien regresi (Goldstein, 1999). Selanjutnya dilakukan perhitungan iterasi menggunakan Iterasi Newton Raphson.

$$\boldsymbol{\beta}_{(m+1)} = \boldsymbol{\beta}_{(m)} - \mathbf{H}^{-1}(\boldsymbol{\beta}_{(m)}) \mathbf{g}(\boldsymbol{\beta}_{(m)}) \quad (2.31)$$

Iterasi berhenti jika

$$\|\boldsymbol{\beta}_{(m+1)} - \boldsymbol{\beta}_{(m)}\| < \varepsilon$$

ε bilangan positif yang sangat kecil.

Langkah-langkah iterasi Newton Raphson diberikan sebagai berikut,

- Menentukan nilai dugaan awal $\boldsymbol{\beta}_{(0)}$ kemudian menghitung $\pi(\mathbf{X}_i)_{(0)}$.
- Dari $\pi(\mathbf{X}_i)_{(0)}$ pada langkah a. diperoleh matriks Hessian $\mathbf{H}(\boldsymbol{\beta})$ vektor gradien $\mathbf{g}(\boldsymbol{\beta})$.
- Kemudian menghitung $\boldsymbol{\beta}_{(0)}$ dengan persamaan $\boldsymbol{\beta}_{m+1} = \boldsymbol{\beta}_m - \mathbf{H}^{-1}(\boldsymbol{\beta}_m) \mathbf{g}(\boldsymbol{\beta}_m)$, dengan $m=0$.
- Menghitung $\|\boldsymbol{\beta}_{m+1} - \boldsymbol{\beta}_m\| < \varepsilon$, kemudian ulangi langkah b sampai d.
- Iterasi berhenti hingga nilai $\boldsymbol{\beta}$ konvergen atau $\|\boldsymbol{\beta}_{m+1} - \boldsymbol{\beta}_m\| < \varepsilon$.

2.7 Pengujian Signifikansi Parameter

Pengujian signifikansi parameter dilakukan secara serentak dan parsial. Pengujian serentak dilakukan untuk mengetahui signifikansi parameter β terhadap variabel dependen secara keseluruhan. Hipotesis yang digunakan sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{Paling sedikit terdapat satu } \beta_i \neq 0, i = 1, 2, \dots, p$$

Statistik uji :

$$G^2 = -2[\ln L(\hat{w}) - \ln L(\hat{\Omega})] \quad (2.32)$$

Kriteria penolakan H_0 jika $G^2 > \chi^2_{(k)}$.

Setelah dilakukan pengujian signifikansi parameter serentak, selanjutnya dilakukan pengujian parameter secara parsial. Uji parsial digunakan untuk menguji pengaruh setiap β_i secara individual. Hasil pengujian secara individual akan menunjukkan suatu variabel independen layak untuk masuk ke dalam model atau tidak. Hipotesis yang digunakan sebagai berikut.

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \neq 0 \text{ dengan } i = 1, 2, \dots, p$$

Statistik uji :

$$Z = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \quad (2.33)$$

Statistik uji tersebut, mengikuti distribusi normal sehingga H_0 ditolak jika $|Z| > Z_{\alpha/2}$.

2.8 Pemilihan Model Terbaik

Pemilihan model terbaik dilakukan dengan melihat nilai *Deviance*. *Deviance* merupakan suatu ukuran yang dapat digunakan untuk menentukan kesesuaian suatu model. Secara umum, *Deviance* dapat didefinisikan sebagai berikut.

$$D = -2 \log \left(\frac{\lambda_0}{\lambda_1} \right) \quad (2.34)$$

Dimana λ_0 adalah fungsi kemungkinan dibawah H_0 pada saat mencapai konvergen dan λ_1 adalah fungsi kemungkinan dibawah H_1 pada saat mencapai konvergen.

Semakin kecil nilai *Deviance*, maka model tersebut dikatakan semakin sesuai. Akan tetapi, tidak ada ketentuan yang pasti besar ukuran untuk nilai *Deviance*, sehingga untuk mengetahui suatu model sesuai atau tidak harus dibandingkan dengan model lain (Hox, 2002).

2.9 Ketepatan Klasifikasi

Prosedur klasifikasi adalah suatu evaluasi untuk melihat peluang kesalahan klasifikasi (*misclassification*) dari fungsi klasifikasi. Prosedur klasifikasi yang baik ditentukan dengan nilai misklasifikasi yang kecil. Terdapat alat ukur yang dapat digunakan untuk menentukan kesalahan klasifikasi yaitu *Apparent Error Rate* (APER). Nilai APER menyatakan nilai proporsi sampel yang salah diklasifikasikan oleh fungsi klasifikasi. Penentuan kesalahan pengklasifikasian dapat diketahui melalui tabel ketepatan klasifikasi (Johnson & Wichern, 2007). Tabel ketepatan klasifikasi dapat dilihat pada Tabel 2.3 sebagai berikut.

Tabel 2.3 Tabel Ketepatan Klasifikasi Untuk 2 Kategori

| Aktual | | Prediksi | | Total |
|--------|---|-----------------|-----------------|-----------------|
| | | Y | | |
| | | 0 | 1 | |
| Y | 0 | n ₀₀ | n ₀₁ | n _{0.} |
| | 1 | n ₁₀ | n ₁₁ | n _{1.} |
| Total | | n _{.0} | n _{.1} | n |

Dimana :

n_{00} : Jumlah pengamatan dari Y(0) tepat diklasifikasikan sebagai Y(0)

n_{01} : Jumlah pengamatan dari Y(0) salah diklasifikasikan sebagai Y(1)

n_{10} : Jumlah pengamatan dari Y(1) salah diklasifikasikan sebagai Y(0)

n_{11} : Jumlah pengamatan dari Y(1) tepat diklasifikasikan sebagai Y(1)

$n_{0.}$: Jumlah pengamatan aktual dari Y(0)

$n_{1.}$: Jumlah pengamatan aktual dari Y(1)

$n_{.0}$: Jumlah pengamatan prediksi dari Y(0)

$n_{.1}$: Jumlah pengamatan prediksi dari Y(1)

n : Total pengamatan keseluruhan

Sehingga, dapat dihitung nilai ketepatan klasifikasi dengan rumus sebagai berikut.

$$1 - APER = \frac{\text{Jumlah total objek yang tepat klasifikasi}}{\text{Jumlah total pengamatan}} = \frac{n_{00} + n_{11}}{n} \quad (2.35)$$

Untuk melihat keakuratan ketepatan klasifikasi dari masing-masing metode analisis, dilakukan perhitungan sensitivitas yang digunakan untuk mengetahui besarnya pengamatan yang diklasifikasikan secara tepat. Tabel sensitivitas untuk 2 kategori dapat dilihat pada Tabel 2.4 sebagai berikut.

Tabel 2.4 Tabel Sensitivitas Untuk 2 Kategori

| Tabel 2.11. Tabel Deskripsi Skala 2 Halogen | | | | |
|---------------------------------------------|---|----------|----|-------|
| Aktual | | Prediksi | | Total |
| | | Y | | |
| | | 0 | 1 | |
| Y | 0 | TN | FN | N |
| | 1 | FP | TP | P |
| Total | | N' | P' | N+P |

Dimana :

TN : True Negative (pengamatan dari Y(0) tepat diklasifikasikan sebagai Y(0))

FN : False Negative (pengamatan dari Y(0) salah diklasifikasikan sebagai Y(1))

FP : False Positive (pengamatan dari Y(1) salah diklasifikasikan sebagai Y(0))

TP : True Positive (pengamatan dari Y(1) tepat diklasifikasikan sebagai Y(1))

N : Negative (Jumlah pengamatan yang salah diklasifikasikan)

P : Positive (Jumlah pengamatan yang tepat diklasifikasikan)

Sehingga, dapat dihitung nilai sensitivitas dengan rumus sebagai berikut.

$$\text{Sensitivitas} = \frac{\text{True Positive}}{\text{Positive}} = \frac{TP}{P} \quad (2.36)$$

2.10 Tenaga Kerja Wanita

Pertumbuhan ekonomi yang mendesak serta kurangnya penghasilan, dengan adanya kesempatan untuk bekerja dibidang industri telah memberikan daya tarik bagi tenaga kerja wanita. Banyak alasan yang mendasari bahwa wanita harus bekerja untuk memenuhi kebutuhan ekonomi keluarga karena tidak ada anggota keluarga lain yang mencari nafkah, ada pula yang bekerja karena memang ingin mencapai karier yang baik dalam dunia kerja. Menurut Undang-Undang Nomor 13 Tahun 2003 tentang Ketenagakerjaan diartikan bahwa pekerja wanita adalah wanita

yang sudah dewasa, wanita yang dianggap dewasa disini adalah wanita yang sudah berusia 18 (delapan belas) tahun atau lebih.

Ada banyak faktor dari wanita yang memutuskan untuk bekerja, diantaranya sebagai berikut.

1. Tingkat Pendidikan

Semakin meningkatnya kualitas pendidikan dapat mengakibatkan semakin meningkatnya persaingan dalam mencari pekerjaan. Kondisi ini akan mengakibatkan pemilihan tenaga kerja yang lebih berkualitas dan terjadi pergeseran ke arah pembangunan tenaga kerja dengan pendidikan yang lebih baik untuk jenis pekerjaan yang sama (Sinungan, 2003).

Tingkat pendidikan yang tinggi juga mempengaruhi pendapatan yang diterima oleh seseorang, semakin tinggi tingkat pendidikan dan pengalaman maka makin tinggi pula tingkat pendapatannya. Pada umumnya masyarakat selalu mencari tingkat pendapatan tinggi untuk memenuhi kebutuhan rumah tangganya, akan tetapi dibatasi oleh tingkat pendidikan yang rendah (Nazir, 2010). Selain itu, dengan tingkat pendidikan yang tinggi akan berpeluang untuk memiliki jabatan yang tinggi dalam suatu perusahaan.

2. Usia

Salah satu faktor dari wanita yang memutuskan untuk bekerja adalah usia. Simanjuntak (1998), menyatakan bahwa usia akan mempengaruhi penyediaan tenaga kerja. Penambahan penyediaan tenaga kerja akan mengalami peningkatan sesuai dengan pertambahan usia, kemudian menurun kembali menjelang usia pensiun atau usia tua. Hal ini dikarenakan semakin tinggi tingkat usia maka akan semakin kecil proporsi penduduk yang bersekolah, sehingga penyediaan tenaga kerja mengalami peningkatan.

Ketika semakin tua usia seseorang, tanggung jawab pada keluarga akan semakin besar, terutama pada penduduk usia muda yang sudah menikah. Bagi seseorang yang telah menikah adanya tanggung jawab untuk memenuhi kebutuhan hidup keluarga. Selanjutnya, ketika tingkat usia semakin tua maka akan masuk pada masa pensiun atau yang secara fisik sudah tidak mampu untuk bekerja.

3. Pengeluaran Rumah Tangga

Kegiatan konsumsi tentu dilakukan dalam suatu rumah tangga yang diperlihatkan dari besarnya pengeluaran untuk memenuhi kebutuhan rumah tangga sehari-hari. Pengeluaran rumah tangga diantaranya mencakup berbagai pengeluaran konsumsi akhir rumah tangga atas barang dan jasa dalam memenuhi kebutuhan individu ataupun keluarga. Pengeluaran rumah tangga terdiri atas pengeluaran untuk makanan, pendidikan, kesehatan, perumahan, tabungan, rekreasi, dan lain-lain. Setiap rumah tangga atau kelompok rumah tangga memiliki pola atau struktur konsumsi dan pengeluaran yang berbeda-beda (Rachman dan Wahida, 1998).

Menurut Kaufman dan Hotchkiss (2000), rumah tangga akan cenderung meningkatkan kualitas standar hidup keluarganya. Sehingga, keluarga dengan dua sumber pendapatan (suami dan istri bekerja) akan dirasa lebih dapat meningkatkan kualitas standar hidupnya dibandingkan pada keluarga yang hanya suaminya yang bekerja. Semakin besar pengeluaran rumah tangga per bulan akan memotivasi pekerja untuk meluangkan lebih banyak waktu di dunia kerja. Peningkatan jumlah jam kerja ini diharapkan dapat menutupi pengeluaran rumah tangga yang semakin tinggi.

(Halaman ini sengaja dikosongkan)

BAB 3

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder hasil survei dari *Indonesia Family Life Survey* (IFLS) gelombang 5 pada tahun 2014 atau disebut sebagai IFLS-5 yang dilakukan oleh RAND Labor and Population. Dari survei terhadap keluarga (*Household Survey*) sebanyak 16.204 keluarga telah diwawancara yang diambil 13 dari 27 provinsi di Indonesia, diantaranya yaitu : DKI Jakarta, Jawa Barat, Jawa Timur, Kalimantan Selatan, Sulawesi Selatan, Sumatera Selatan, Nusa Tenggara Barat, Jawa Tengah, Yogyakarta, Bali, Sumatera Utara, Sumatera Barat dan Lampung. Survei juga dilakukan terhadap lingkungan sosial beserta fasilitasnya (*Community-Facility Survey*) yaitu sebanyak 312 lingkungan sosial beserta fasilitasnya. Dalam survei ini tentu saja lingkungan sosial yang dipilih adalah lingkungan sosial yang bersesuaian dengan keluarga yang diwawancara. (Strauss, et. al., 2004).

3.2 Variabel Penelitian

Survei dilakukan pada rumah tangga dan komunitas-fasilitas (Kecamatan), sehingga individu dalam rumah tangga sebagai level 1 terpilih variabel status bekerja ibu rumah tangga (dimana, 1 untuk ibu rumah tangga bekerja dan 0 untuk ibu rumah tangga tidak bekerja) sebagai variabel dependen (Y), variabel pendidikan terakhir, variabel usia dan variabel pengeluaran rumah tangga sebagai variabel independen (X) atau efek tetap (*fixed effect*). Pabrik dalam komunitas-fasilitas (Kecamatan) sebagai level 2 terpilih variabel rata-rata gaji tenaga kerja wanita sebagai variabel independen (Z).

Data seluruhnya berjumlah 4431 wanita yang memiliki status menikah atau pernah menikah. Sebanyak 99 data rata-rata gaji wanita yang bersesuaian dengan data dalam Kecamatan. Secara ringkas, variabel-variabel penelitian yang digunakan dalam penelitian ini dapat dilihat pada Tabel 3.1 sebagai berikut.

Tabel 3.1 Variabel Penelitian

| Variabel | Indikator | Keterangan | Skala |
|----------------------------|---------------------------------|-------------------------------------------------------------------------------------------------------------|---------|
| Y | Status Bekerja Ibu Rumah Tangga | 0: Tidak Bekerja 1: Bekerja | Nominal |
| Level 1 (Individu) | | | |
| X ₁ | Pendidikan Terakhir | 0 : Tidak sekolah 1 : Lulus SD Sederajat 2 : Lulus SMP 3 : Lulus SMA 4 : Lulus Perguruan Tinggi | Ordinal |
| X ₂ | Usia | - | Rasio |
| X ₃ | Pengeluaran Rumah Tangga | - | Rasio |
| Level 2 (Kecamatan) | | | |
| Z | Gaji Tenaga Kerja Wanita | - | Rasio |

Berdasarkan variabel yang digunakan dalam penelitian ini, maka didapatkan definisi operasional variabel sebagai berikut.

1. Status bekerja ibu rumah tangga (Variabel Dependen/Y)
Status bekerja ibu rumah tangga menjelaskan tentang istri dari kepala rumah tangga bekerja atau tidak, dimana terdapat kategori 0 untuk ibu rumah tangga yang tidak bekerja dan 1 untuk ibu rumah tangga yang bekerja.
2. Pendidikan terakhir (Variabel Independen/X₁)
Pendidikan terakhir menjelaskan tentang pendidikan yang telah ditempuh oleh ibu rumah tangga.
3. Usia (Variabel Independen/X₂)
Usia menjelaskan tentang usia dari wanita yang berstatus menikah saat tahun 2014 yang dinyatakan dalam satuan tahun (th).
4. Pengeluaran rumah tangga (Variabel Independen/X₃)
Pengeluaran rumah tangga menjelaskan tentang seberapa banyak pengeluaran yang telah dikeluarkan suatu rumah tangga selama sebulan, meliputi pengeluaran untuk makanan dan non makanan yang dinyatakan dalam satuan Rupiah (Rp).
5. Rata-rata gaji tenaga kerja wanita (Variabel Independen/Z)
Rata-rata gaji tenaga kerja wanita menjelaskan tentang gaji yang diberikan oleh pabrik untuk tenaga kerja wanita selama sebulan yang dinyatakan dalam satuan Rupiah (Rp).

3.3 Struktur Data Penelitian

Struktur data dalam penelitian ini merupakan data hirarki, yaitu data individu yang terkelompok di dalam Kecamatan. Sehingga, struktur data dalam penelitian ini dapat dilihat pada Tabel 3.2 sebagai berikut.

Tabel 3.2 Struktur Data Model 2 Level

| Kecamatan | Individu | Variabel Dependen | Variabel Independen Level 1 | | | Variabel Independen Level 2 |
|-----------|----------|----------------------|--------------------------------|----------------------|----------------------|--------------------------------|
| | | Y | X ₁ | X ₂ | X ₃ | Z ₁ |
| 1 | 1 | Y ₁₁ | X ₁₁₁ | X ₂₁₁ | X ₃₁₁ | Z ₁₁ |
| | 2 | Y ₂₁ | X ₁₂₁ | X ₂₂₁ | X ₃₂₁ | |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | |
| | 56 | Y _{56;1} | X _{1;56;1} | X _{2;56;1} | X _{3;56;1} | |
| 2 | 1 | Y ₁₂ | X ₁₁₂ | X ₂₁₂ | X ₃₁₂ | Z ₁₂ |
| | 2 | Y ₂₂ | X ₁₂₂ | X ₂₂₂ | X ₃₂₂ | |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | |
| | 41 | Y _{41;2} | X _{1;41;2} | X _{2;41;2} | X _{3;41;2} | |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 99 | 1 | Y _{1;99} | X _{1;1;99} | X _{2;1;99} | X _{3;1;99} | Z _{1;99} |
| | 2 | Y _{2;99} | X _{1;2;99} | X _{2;2;99} | X _{3;2;99} | |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | |
| | 68 | Y _{68;99} | X _{1;68;99} | X _{2;68;99} | X _{3;68;99} | |

3.4 Langkah-Langkah Analisis

Adapun langkah-langkah yang dilakukan untuk menganalisis penelitian tentang faktor yang mempengaruhi status bekerja ibu rumah tangga di Indonesia menggunakan model multilevel dengan respon biner sebagai berikut.

1. Analisis statistika deskriptif dari karakteristik status bekerja ibu rumah tangga di Indonesia.
2. Mengestimasi parameter dalam model 2 level dengan respon biner menggunakan MLE dengan pendekatan prosedur Gaussian-Hermite Quadrature dan iterasi Newton Raphson.
3. Membuat model menggunakan analisis regresi logistik tanpa melibatkan efek Kecamatan dengan langkah-langkah sebagai berikut.
 - a. Melakukan pengujian signifikansi parameter baik secara serentak maupun parsial pada koefisien regresi dalam model yang dihasilkan.
 - b. Membuat model regresi logistik biner.

$$\ln\left(\frac{\pi(X)}{1-\pi(X)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon_i$$

- c. Menginterpretasikan model regresi logistik biner dengan variabel yang signifikan.
 - d. Menentukan ukuran kesesuaian model.
4. Membuat model menggunakan model 2 level dengan respon biner melibatkan efek Kecamatan untuk langkah-langkah sebagai berikut.
- a. Variabel pada level 1 (individu) adalah variabel status bekerja ibu rumah tangga (Y) dan 3 variabel independen yaitu variabel pendidikan terakhir (X_1), variabel usia (X_2) dan variabel pengeluaran rumah tangga (X_3). Sehingga, model pada level 1 sebagai berikut.

$$\ln\left(\frac{\pi(X_{ij})}{1-\pi(X_{ij})}\right) = \beta_{0j} + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i ; i = 1, 2, \dots, n_m ; j = 1, 2, \dots, m$$

- b. Menerapkan model level 1 pada kasus status bekerja ibu rumah tangga dengan menggunakan model regresi logistik biner untuk masing-masing Kecamatan yaitu sebanyak 99 Kecamatan dengan langkah-langkah sebagai berikut.

- Melakukan pengujian signifikansi parameter yang dihasilkan pada model level 1 baik secara serentak maupun parsial.
- Membuat model pada level 1 dengan model regresi logistik biner. Sehingga, model pada level 1 sebagai berikut.

$$\ln\left(\frac{\pi(X_{i1})}{1-\pi(X_{i1})}\right) = \beta_{01} + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$$

$$\ln\left(\frac{\pi(X_{i2})}{1-\pi(X_{i2})}\right) = \beta_{02} + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$$

⋮

$$\ln\left(\frac{\pi(X_{i;99})}{1-\pi(X_{i;99})}\right) = \beta_{0;99} + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$$

- Menginterpretasikan model regresi logistik biner untuk variabel yang signifikan pada masing-masing Kecamatan dalam model level 1.

- c. Variabel pada level 2 (Kecamatan) adalah random intercept yaitu β_0 sebagai variabel respon. Sehingga, model pada level 2 tanpa melibatkan variabel Z ke dalam model sebagai berikut.

$$\beta_{0j} = \beta_{00} + u_{0j}$$

- d. Menerapkan model level 2 pada kasus status bekerja ibu rumah tangga dengan menggunakan model regresi sederhana dengan langkah-langkah sebagai berikut.

- Melakukan pengujian signifikansi parameter yang dihasilkan pada model level 2 baik secara serentak maupun parsial.
- Membuat model pada level 2 dengan dengan intersep acak, yaitu tidak melibatkan variabel Z ke dalam model. Sehingga, model pada level 2 sebagai berikut.

$$\beta_{0j} = \beta_{00} + u_{0j}$$

- Menggabungkan model pada level 1 dan model pada level 2 dengan intersep acak tanpa melibatkan variabel Z ke dalam model.

$$\ln \left(\frac{\pi(X_{ij})}{1 - \pi(X_{ij})} \right) = \beta_{00} + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + u_{0j} + \varepsilon_i$$

- Menginterpretasikan model 2 level dengan respon biner untuk variabel yang signifikan.
- Menentukan ukuran kesesuaian model.

- e. Variabel pada level 2 (Kecamatan) adalah random intercept yaitu β_0 sebagai variabel respon dan 1 variabel independen yaitu variabel rata-rata gaji wanita (Z). Sehingga, model pada level 2 dengan melibatkan variabel Z ke dalam model sebagai berikut.

$$\beta_{0j} = \beta_{00} + \beta_{01} Z_j + u_{0j}$$

- f. Menerapkan model level 2 pada kasus status bekerja ibu rumah tangga dengan menggunakan model regresi sederhana dengan langkah-langkah sebagai berikut.

- Melakukan pengujian signifikansi parameter yang dihasilkan pada model level 2 baik secara serentak maupun parsial.

- Membuat model level 2 dengan dengan intersep acak, yaitu melibatkan variabel Z ke dalam model. Sehingga, model pada level 2 sebagai berikut.

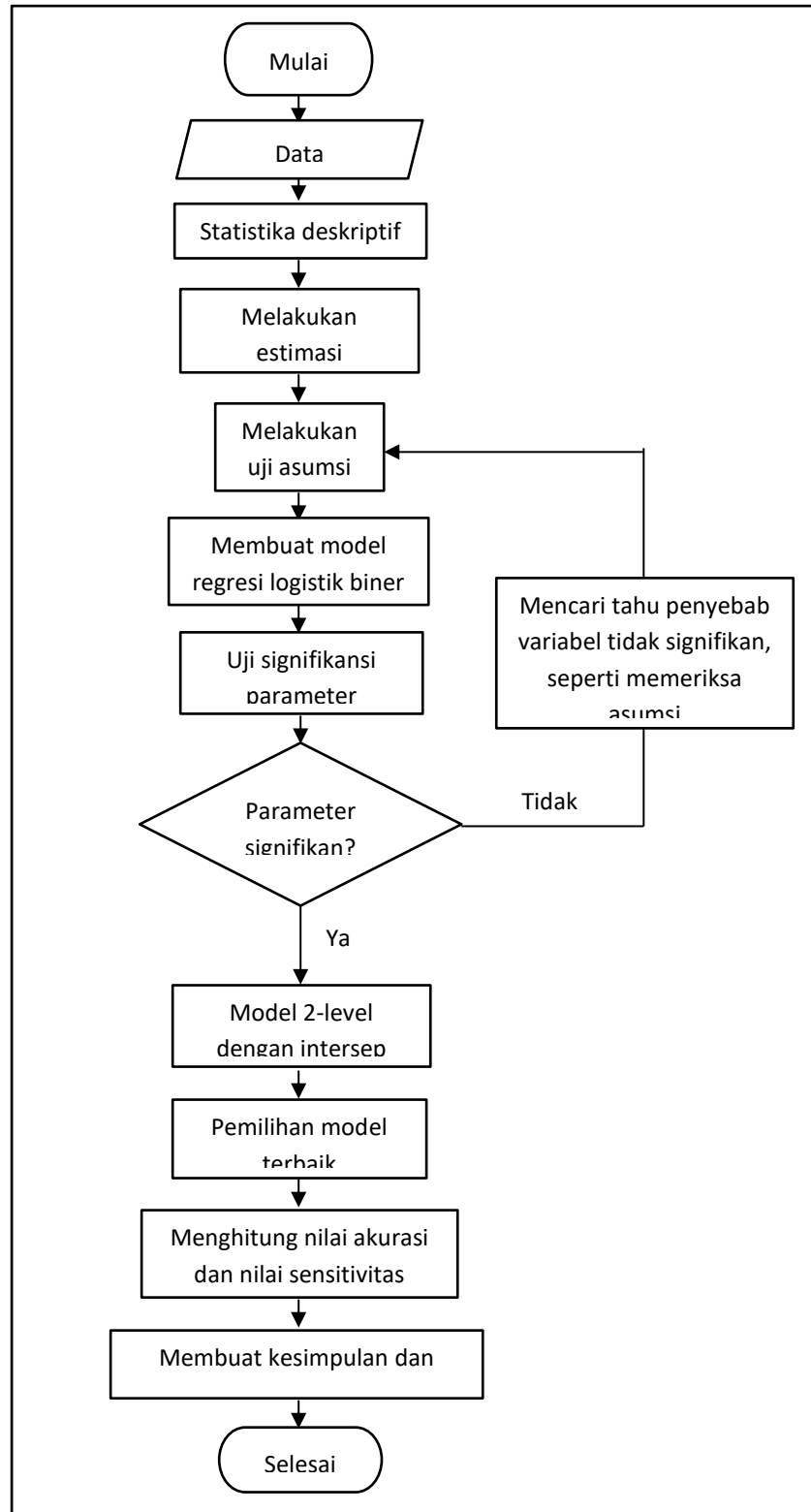
$$\beta_{0j} = \beta_{00} + \beta_{01}Z_j + u_{0j}$$

- Menggabungkan model level 1 dan model level 2 dengan intersep acak dengan melibatkan variabel Z ke dalam model.

$$\ln \left(\frac{\pi(X_{ij})}{1 - \pi(X_{ij})} \right) = \beta_{00} + \beta_{01}Z_j + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + u_{0j} + \varepsilon_i$$

- Menginterpretasikan model 2 level dengan respon biner untuk variabel yang signifikan.
 - Menentukan ukuran kesesuaian model.
5. Melakukan pemilihan model terbaik dengan cara membandingkan ukuran kesesuaian antar model yang dihasilkan.
 6. Menghitung ketepatan klasifikasi dan sensitivitas pada masing-masing model yang dihasilkan.
 7. Membuat kesimpulan dan saran.

Diagram alir (*flow chart*) dalam penelitian ini dapat digambarkan seperti pada Gambar 3.1 sebagai berikut.



Gambar 3.1 Diagram alir penelitian

(Halaman ini sengaja dikosongkan)

BAB 4

ANALISIS DAN PEMBAHASAN

Analisis yang dilakukan dalam penelitian ini meliputi estimasi parameter, analisis karakteristik data, pengujian asumsi, analisis regresi logistik biner untuk data pada level 1 dan analisis regresi multilevel dengan respon biner pada data gabungan antara level 1 dan level 2. Hasil analisis dan pembahasan dapat dilihat sebagai berikut.

4.1 Estimasi Parameter Model Regresi Multilevel Respon Biner

Metode yang digunakan untuk menentukan kesesuaian model dalam model regresi multilevel yaitu *Generalized Linear Mixed Model* (GLMM) (Agresti, 2007). Secara umum, model regresi 2-level dengan *random intercept* sebagai berikut

$$Y_{ij} = \beta X_{ij} + u_j + \varepsilon_i \quad (4.1)$$

Dengan u_j berdistribusi normal dengan rata-rata 0 dan varians σ_u^2 , dan u_j adalah error untuk level-2 yang independen terhadap ε_i .

Variabel Y dalam model multilevel linier berdistribusi normal yang merupakan kombinasi linier dari koefisien acak. Model multilevel non linier atau lebih khususnya adalah model dengan respon biner merupakan model regresi yang tidak bisa dijelaskan sebagai kombinasi linier dari koefisiennya sehingga dengan demikian variabel Y tidak lagi berdistribusi normal, tetapi berdistribusi binomial.

Data variabel respon yang digunakan dalam model multilevel dalam penelitian ini adalah data dengan skala nominal dengan hanya berupa 2 kategori yaitu “sukses” atau “gagal” yang dinotasikan dengan $Y=1$ (sukses) dan $Y=0$ (gagal).

Variabel Y mengikuti distribusi Bernoulli untuk setiap observasi tunggal, sehingga fungsi probabilitas untuk setiap observasi diberikan sebagai berikut.

$$f(Y_i) = (\pi(X_i))^{Y_i} (1 - \pi(X_i))^{1-Y_i} \quad ; Y_i = 0,1 \quad (4.2)$$

Untuk model dengan fungsi penghubung logit model (4.1) dapat dituliskan dalam bentuk matriks sebagai berikut.

$$\text{Logit}(\pi_{ij}) = \mathbf{X}_{ij}\boldsymbol{\beta} + \mathbf{u}_j \quad (4.3)$$

Dengan \mathbf{X}_{ij} adalah vektor variabel independen, $\boldsymbol{\beta}$ adalah vektor koefisien regresi dan \mathbf{u}_j adalah vektor error untuk level 2.

Misalkan diambil n sampel acak dalam suatu populasi ($Y_{ij}, X_{1ij}, X_{2ij}, \dots, X_{kij}$; $i = 1, 2, \dots, n$; $j = 1, 2, \dots, m$) dimana $Y_{ij} \sim B(1, \pi(\mathbf{X}_{ij}))$ adalah independen dan $\mathbf{X}_{ij} = [1 \ X_{1ij} \ X_{2ij} \ \dots \ X_{kij}]^T$. Selanjutnya adalah membuat fungsi likelihood dari Y sebagai berikut.

$$\begin{aligned} L &= f(Y_{1ij}, Y_{2ij}, \dots, Y_{nij}) \\ L &= f(Y_{1ij}), f(Y_{2ij}), \dots, f(Y_{nij}) \\ L &= \prod_{i=1}^n f(Y_{ij}) \\ L &= \prod_{i=1}^n (\pi(\mathbf{X}_{ij}))^{Y_{ij}} (1 - \pi(\mathbf{X}_{ij}))^{1-Y_{ij}} \\ L &= \prod_{i=1}^n (\pi(\boldsymbol{\beta}\mathbf{X}_{ij} + \mathbf{u}_j))^{Y_{ij}} (1 - \pi(\boldsymbol{\beta}\mathbf{X}_{ij} + \mathbf{u}_j))^{1-Y_{ij}} \end{aligned} \quad (4.4)$$

Dari persamaan (4.1), dapat diketahui bahwa u_j berdistribusi normal dengan rata-rata 0 dan varians σ_u^2 , dan \mathbf{u}_j adalah *random effect* untuk level-2 sehingga fungsi likelihood harus diintegrasikan yang nantinya digunakan untuk pendekatan integrasi sebagai berikut.

$$\begin{aligned} L &= \int_{-\infty}^{\infty} \prod_{i=1}^n f(Y_{ij}) f(\mathbf{u}_j) d\mathbf{u}_j \\ L &= \int_{-\infty}^{\infty} \left[\prod_{i=1}^n (\boldsymbol{\beta}\mathbf{X}_{ij} + \mathbf{u}_j)^{Y_{ij}} (1 - (\boldsymbol{\beta}\mathbf{X}_{ij} + \mathbf{u}_j))^{1-Y_{ij}} \right] \frac{1}{\sqrt{2\pi\sigma_u^2}} \exp\left(-\frac{1}{2\sigma_u^2} \mathbf{u}_j^2\right) d\mathbf{u}_j \\ L &= \int_{-\infty}^{\infty} \left[\prod_{i=1}^n f(Y_{ij} | \boldsymbol{\beta}, \mathbf{X}_{ij}, \mathbf{u}_j) \right] \frac{1}{\sqrt{2\pi\sigma_u^2}} \exp\left(-\frac{1}{2\sigma_u^2} \mathbf{u}_j^2\right) d\mathbf{u}_j \end{aligned} \quad (4.5)$$

Misalkan $\mathbf{u} = \sqrt{2\sigma_u^2} v = \theta v$, dimana $v = \frac{\mathbf{u}}{\sqrt{2\sigma_u^2}}$ dan diberikan

$$g(v) = \prod_{i=1}^n f(Y_{ij} | \boldsymbol{\beta}, \mathbf{X}_{ij}, \mathbf{u}_j)$$

sehingga diperoleh fungsi likelihood sebagai berikut.

$$L = \int_{-\infty}^{\infty} (g(v)) \frac{1}{\theta\sqrt{\pi}} \exp(-v^2) \theta dv$$

$$L = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (g(v)) \exp(-v^2) dv \quad (4.6)$$

Dari persamaan (4.6), dapat diselesaikan dengan menggunakan metode numerik yaitu pendekatan integral kuadrat Gauss-Hermite Quadrature sebagai berikut.

$$\int_{-\infty}^{\infty} (g(v)) \exp(-v^2) dv \approx \sum_{k=1}^K w_k g(v_k)$$

dimana :

w_k : Bobot quadrature

v_k : Titik quadrature

Setelah diperoleh pendekatan integral kuadrat Gauss-Hermite Quadrature, fungsi likelihood menjadi

$$L = \frac{1}{\sqrt{\pi}} \sum_{k=1}^K w_k \left[\prod_{i=1}^n f(Y_{ij} | X_{ij}, \beta, \sqrt{2\sigma_u^2} v_k) \right]$$

$$L = \prod_{j=1}^m \frac{1}{\sqrt{\pi}} \sum_{k=1}^K w_k \left[\prod_{i=1}^n f(Y_{ij} | X_{ij}, \beta, \sqrt{2\sigma_u^2} v_k) \right] \quad (4.7)$$

Diketahui bahwa $\sigma_u^2 = \frac{\rho}{(1-\rho)}$ sehingga $\sigma_u^2 = \left(\frac{\rho}{(1-\rho)} \right)^{\frac{1}{2}}$ (Greene, 2002), maka

fungsi likelihood pada persamaan (4.4) dan (4.7) menjadi

$$L(\beta) = \prod_{j=1}^m \frac{1}{\sqrt{\pi}} \sum_{k=1}^K w_k \left[\prod_{i=1}^n \left(X_{ij} \beta + \sqrt{2\sigma_u^2} v_k \right)^{Y_{ij}} \left(1 - \left(X_{ij} \beta + \sqrt{2\sigma_u^2} v_k \right) \right)^{1-Y_{ij}} \right]$$

$$L(\beta) = \prod_{j=1}^m \frac{1}{\sqrt{\pi}} \sum_{k=1}^K w_k \left[\prod_{i=1}^n \left(X_{ij} \beta + \sqrt{\frac{2\rho}{(1-\rho)}} v_k \right)^{Y_{ij}} \left(1 - \left(X_{ij} \beta + \sqrt{\frac{2\rho}{(1-\rho)}} v_k \right) \right)^{1-Y_{ij}} \right]$$

$$L(\beta) = \prod_{j=1}^m \frac{1}{\sqrt{\pi}} \sum_{k=1}^K w_k \left[\prod_{i=1}^n \left(X_{ij} \beta + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{Y_{ij}} \left(1 - \left(X_{ij} \beta + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right) \right)^{1-Y_{ij}} \right]$$

Langkah selanjutnya adalah membuat \ln dari fungsi likelihood untuk memaksimumkan fungsi likelihood sebagai berikut.

$$\ln L = \left[\prod_{j=1}^m \frac{1}{\sqrt{\pi}} \sum_{k=1}^K w_k \left[\prod_{i=1}^n \left(X_{ij} \beta + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{Y_{ij}} \left(1 - \left(X_{ij} \beta + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right) \right)^{1-Y_{ij}} \right] \right]$$

$$\begin{aligned}
\ln L &= \ln \frac{1}{\sqrt{\pi}} + \ln \left[\prod_{j=1}^m \sum_{k=1}^K w_k \left[\prod_{i=1}^n \left(X_{ij} \boldsymbol{\beta} + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{Y_{ij}} \left(1 - \left(X_{ij} \boldsymbol{\beta} + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{1-Y_{ij}} \right) \right] \right] \\
\ln L &= \ln \frac{1}{\sqrt{\pi}} + \ln \left[\prod_{j=1}^m \sum_{k=1}^K \prod_{i=1}^n w_k \left[\left(X_{ij} \boldsymbol{\beta} + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{Y_{ij}} \left(1 - \left(X_{ij} \boldsymbol{\beta} + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{1-Y_{ij}} \right) \right] \right] \\
\ln L &= \ln \frac{1}{\sqrt{\pi}} + \ln \left[\prod_{j=1}^m \sum_{k=1}^K \prod_{i=1}^n w_k \left[\left(X_{ij} \boldsymbol{\beta} + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{Y_{ij}} \right] + \left[1 - \left(\prod_{j=1}^m \sum_{k=1}^K \prod_{i=1}^n w_k \left(X_{ij} \boldsymbol{\beta} + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{1-Y_{ij}} \right) \right] \right] \\
\ln L &= \ln \frac{1}{\sqrt{\pi}} + \sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k \left(X_{ij} \boldsymbol{\beta} + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{Y_{ij}} + \left(1 - \left(\sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k \left(X_{ij} \boldsymbol{\beta} + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{1-Y_{ij}} \right) \right)
\end{aligned}$$

Kemudian memaksimumkan $\ln L$ dengan menurunkan fungsi \ln dari fungsi likelihood terhadap $\boldsymbol{\beta}$ sebagai berikut.

- Turunan Pertama

$$\begin{aligned}
\frac{\partial \ln L}{\partial \boldsymbol{\beta}^T} &= \begin{bmatrix} \frac{\partial \ln L}{\partial \beta_0} \\ \frac{\partial \ln L}{\partial \beta_1} \\ \frac{\partial \ln L}{\partial \beta_2} \\ \vdots \\ \frac{\partial \ln L}{\partial \beta_k} \end{bmatrix} = \mathbf{0} = \mathbf{g}(\boldsymbol{\beta}) \\
\mathbf{g}(\boldsymbol{\beta}) &= \frac{\partial \left(\ln \frac{1}{\sqrt{\pi}} + \sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k \left(X_{ij} \boldsymbol{\beta} + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{Y_{ij}} + \left(1 - \left(\sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k \left(X_{ij} \boldsymbol{\beta} + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{1-Y_{ij}} \right) \right) \right)}{\partial \boldsymbol{\beta}^T} \\
\mathbf{g}(\boldsymbol{\beta}) &= \left(\sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k Y_{ij} \left(X_{ij}^T \boldsymbol{\beta} + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{Y_{ij}-1} (X_{ij}^T) + \left(1 - \left(\sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k (1-Y_{ij}) \left(X_{ij}^T \boldsymbol{\beta} + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{-Y_{ij}} \right) \right) (-X_{ij}^T) \right) \\
\mathbf{g}(\boldsymbol{\beta}) &= \mathbf{0} \\
\sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k X_{ij}^T Y_{ij} \left(X_{ij}^T \boldsymbol{\beta} + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{Y_{ij}-1} + \left(1 + \left(\sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k X_{ij}^T (1-Y_{ij}) \left(X_{ij}^T \boldsymbol{\beta} + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{-Y_{ij}} \right) \right) &= \mathbf{0}
\end{aligned}$$

- Turunan Kedua

$$H(\beta) = \frac{\partial^2 \ln L}{\partial \beta \partial \beta^T}$$

$$H(\beta) = \frac{\partial \left(\frac{\partial \ln L}{\partial \beta^T} \right)}{\partial \beta}$$

$$\frac{\partial \ln L}{\partial \beta^T} = \sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k \mathbf{X}_{ij}^T Y_{ij} \left(\mathbf{X}_{ij}^T \beta + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{Y_{ij}-1} + \left(1 + \left(\sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k \mathbf{X}_{ij}^T (1-Y_{ij}) \left(\mathbf{X}_{ij}^T \beta + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{-Y_{ij}} \right) \right)$$

$$\frac{\partial \ln L}{\partial \beta^T} = \sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k \mathbf{X}_{ij}^T Y_{ij} (C)^{Y_{ij}-1} + \left(1 + \left(\sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k \mathbf{X}_{ij}^T (1-Y_{ij}) (C)^{-Y_{ij}} \right) \right)$$

$$\frac{\partial C}{\partial \beta^T} = \frac{\partial \left(\mathbf{X}_{ij}^T \beta + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)}{\partial \beta^T} = \mathbf{X}_{ij}$$

$$\frac{\partial^2 \ln L}{\partial \beta \partial \beta^T} = \frac{\sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k \mathbf{X}_{ij}^T Y_{ij} (C)^{Y_{ij}-1} + \left(1 + \left(\sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k \mathbf{X}_{ij}^T (1-Y_{ij}) (C)^{-Y_{ij}} \right) \right)}{\partial \beta}$$

$$\frac{\partial^2 \ln L}{\partial \beta \partial \beta^T} = \sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k \mathbf{X}_{ij}^T Y_{ij} \left((Y_{ij}-1)(C)^{Y_{ij}-2} \mathbf{X}_{ij}^T \right) + \left(1 + \left(\sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k \mathbf{X}_{ij}^T (1-Y_{ij}) \left(-Y_{ij} (C)^{-Y_{ij}-1} - \mathbf{X}_{ij}^T \right) \right) \right)$$

$$\frac{\partial^2 \ln L}{\partial \beta \partial \beta^T} = \sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k \mathbf{X}_{ij}^T Y_{ij} \left((Y_{ij}-1) \left(\mathbf{X}_{ij}^T \beta + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{Y_{ij}-2} \mathbf{X}_{ij}^T \right) + \left(1 + \left(\sum_{j=1}^m \sum_{k=1}^K \sum_{i=1}^n w_k \mathbf{X}_{ij}^T (1-Y_{ij}) \left(-Y_{ij} \left(\mathbf{X}_{ij}^T \beta + \left(\frac{2\rho}{(1-\rho)} \right)^{\frac{1}{2}} v_k \right)^{-Y_{ij}-1} - \mathbf{X}_{ij}^T \right) \right) \right)$$

$$H(\beta) = \begin{bmatrix} \frac{\partial^2 \ln L}{\partial \beta_0^2} & \frac{\partial^2 \ln L}{\partial \beta_0 \partial \beta_1} & \frac{\partial^2 \ln L}{\partial \beta_0 \partial \beta_2} & \dots & \frac{\partial^2 \ln L}{\partial \beta_0 \partial \beta_k} \\ & \frac{\partial^2 \ln L}{\partial \beta_1^2} & \frac{\partial^2 \ln L}{\partial \beta_1 \partial \beta_2} & \dots & \frac{\partial^2 \ln L}{\partial \beta_1 \partial \beta_k} \\ & & \frac{\partial^2 \ln L}{\partial \beta_2^2} & \dots & \frac{\partial^2 \ln L}{\partial \beta_2 \partial \beta_k} \\ & \vdots & & \ddots & \vdots \\ \text{simetris} & & & & \frac{\partial^2 \ln L}{\partial \beta_k^2} \end{bmatrix}$$

Hasil yang diperoleh tidak *close form* karena nilai β berbentuk skalar, sehingga nilai awal β dicari dengan menggunakan rumus GLS (*Generalized Least Square*) sebagai berikut (Goldstein, 1999).

$$\hat{\beta} = (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1} \mathbf{Y}$$

Selanjutnya nilai β dicari menggunakan iterasi numerik yaitu iterasi Newton Raphson dengan rumus sebagai berikut.

$$\beta_{(m+1)} = \beta_{(m)} - H^{-1}(\beta_{(m)})g(\beta_{(m)})$$

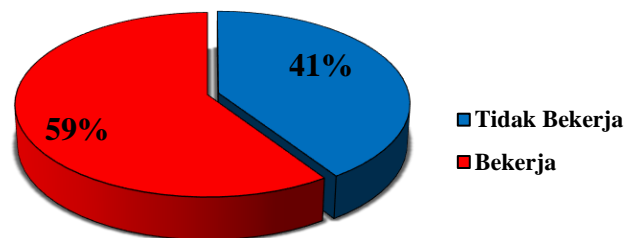
Iterasi berhenti jika $\|\beta_{(m+1)} - \beta_{(m)}\| < \varepsilon$, ε bilangan positif yang sangat kecil.

4.2 Karakteristik Data

Karakteristik data dianalisis dengan menggunakan analisis statistika deskriptif. Langkah pertama adalah melakukan analisis statistika deskriptif untuk mengetahui karakteristik status bekerja IRT. Statistika deskriptif untuk data kategorik menggunakan *pie chart* dan *bar chart*, serta statistika deskriptif untuk data kontinu menggunakan tabel deskriptif yang meliputi jumlah keseluruhan data, rata-rata, standar deviasi, nilai minimum dan nilai maksimum. Data yang digunakan dalam analisis statistika deskriptif adalah data pada Lampiran 1.

4.2.1 Status Bekerja Ibu Rumah Tangga

Statistika deskriptif untuk karakteristik status bekerja ibu rumah tangga disajikan dengan menggunakan *pie chart*. Hasil karakteristik dapat dilihat pada Gambar 4.1 sebagai berikut.



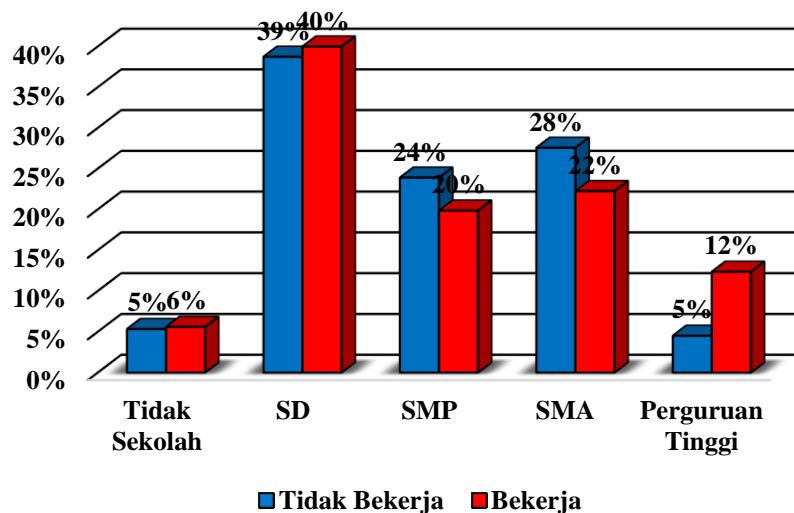
Gambar 4.1 Persentase status bekerja Ibu Rumah Tangga

Berdasarkan Gambar 4.1, dapat diketahui bahwa dari total keseluruhan sampel sebanyak 4431 ibu rumah tangga yang ada di Indonesia, ibu rumah tangga yang tidak bekerja sebesar 41% atau sebanyak 1800 ibu rumah tangga dan ibu rumah tangga yang bekerja sebesar 59% atau sebanyak 2631 ibu rumah tangga. Hal ini menunjukkan bahwa mayoritas ibu rumah tangga yang ada di Indonesia pada tahun 2014 lebih banyak yang memutuskan untuk bekerja dan menjadi ibu rumah

tangga dibandingkan dengan yang memutuskan untuk tidak bekerja atau hanya menjadi ibu rumah tangga.

4.2.2 Pendidikan Terakhir

Statistika deskriptif untuk karakteristik pendidikan terakhir ibu rumah tangga yang berstatus tidak bekerja dan berstatus bekerja disajikan dengan menggunakan *bar chart*. Hasil karakteristik dapat dilihat pada Gambar 4.2 sebagai berikut.



Gambar 4.2 Persentase pendidikan terakhir ibu rumah tangga

Berdasarkan Gambar 4.2, dapat diketahui bahwa dari total keseluruhan sampel sebanyak 1800 ibu rumah tangga yang berstatus tidak bekerja yang ada di Indonesia, yang paling tinggi adalah yang memiliki latar belakang pendidikan terakhir SD yaitu sebesar 39% atau sebanyak 697 ibu rumah tangga dan yang paling rendah adalah yang memiliki latar belakang tidak sekolah sebesar 5% atau sebanyak 96 ibu rumah tangga dan pendidikan terakhir perguruan tinggi yaitu sebesar 5% atau sebanyak 81 ibu rumah tangga. Sedangkan, dari total keseluruhan sampel sebanyak 2631 ibu rumah tangga yang berstatus bekerja yang ada di Indonesia, yang paling tinggi adalah yang memiliki latar belakang pendidikan terakhir SD yaitu sebesar 40% atau sebanyak 1052 ibu rumah tangga dan yang paling rendah adalah yang memiliki latar belakang tidak sekolah yaitu sebesar 6% atau sebanyak 147 ibu rumah tangga. Jika dilihat dari hasil pola *bar chart*, dapat diketahui bahwa

ibu rumah tangga yang memutuskan untuk ikut serta dalam dunia kerja mayoritas adalah ibu rumah tangga yang memiliki latar belakang tidak sekolah, pendidikan terakhir SD dan perguruan tinggi. Ibu rumah tangga dengan latar belakang tidak sekolah dan pendidikan terakhir SD memiliki motivasi untuk bekerja yang lebih tinggi kemungkinan karena adanya tuntutan ekonomi sehingga mengharuskan untuk bekerja agar dapat membantu memenuhi kebutuhan rumah tangga. Sedangkan, ibu rumah tangga dengan latar belakang pendidikan terakhir perguruan tinggi memiliki motivasi untuk bekerja yang lebih tinggi kemungkinan karena ingin menyalurkan keterampilan yang dimiliki dan mengaplikasikannya dalam dunia kerja, serta untuk menjalin hubungan sosial. Selain itu, motivasi ibu rumah tangga yang tidak ikut serta dalam dunia kerja mayoritas adalah ibu rumah tangga yang memiliki latar belakang pendidikan SMP dan SMA. Ibu rumah tangga dengan latar belakang pendidikan SMP dan SMA memiliki motivasi untuk bekerja yang lebih rendah kemungkinan karena kebutuhan rumah tangga sudah cukup terpenuhi tanpa ibu rumah tangga ikut serta dalam dunia kerja.

4.2.3 Usia

Statistika deskriptif untuk karakteristik usia ibu rumah tangga yang berstatus tidak bekerja dan berstatus bekerja disajikan dengan menggunakan tabel statistika deskriptif. Hasil karakteristik dapat dilihat pada Tabel 4.1 sebagai berikut.

Tabel 4.1 Statistika Deskriptif Untuk Usia

| Variabel | n | Rata-rata | Standar Deviasi | Minimum | Maksimum |
|-----------------|----------|------------------|------------------------|----------------|-----------------|
| Tidak Bekerja | 1800 | 41 | 15 | 15 | 94 |
| Bekerja | 2631 | 42 | 12 | 15 | 98 |

Berdasarkan Tabel 4.1, dapat diketahui bahwa dari total keseluruhan sebanyak 1800 ibu rumah tangga yang berstatus tidak bekerja yang ada di Indonesia, rata-rata usia yaitu 41 tahun dan standar deviasi sebesar 15 tahun dengan usia minimum ibu rumah tangga adalah 15 tahun dan usia maksimum ibu rumah tangga adalah 94 tahun. Sedangkan, dari total keseluruhan sebanyak 2631 ibu rumah tangga yang berstatus bekerja yang ada di Indonesia, rata-rata usia yaitu 42 dan standar deviasi sebesar 12 tahun dengan usia minimum ibu rumah tangga adalah 15 tahun dan usia maksimum ibu rumah tangga adalah 98 tahun. Hal ini menunjukkan bahwa rentang

usia antara ibu rumah tangga yang berstatus tidak bekerja dan berstatus bekerja tidak jauh berbeda, yaitu antara usia 26 tahun sampai dengan umur 56 tahun.

4.2.4 Pengeluaran Rumah Tangga

Statistika deskriptif untuk karakteristik pengeluaran rumah tangga dari ibu rumah tangga yang berstatus tidak bekerja dan berstatus bekerja disajikan dengan menggunakan tabel statistika deskriptif. Hasil karakteristik dapat dilihat pada Tabel 4.2 sebagai berikut.

Tabel 4.2 Statistika Deskriptif Untuk Pengeluaran Rumah Tangga

| Variabel | n | Rata-rata | Standar Deviasi | Minimum | Maksimum |
|-----------------|----------|------------------|------------------------|----------------|-----------------|
| Tidak Bekerja | 1800 | 1.456.451 | 2.735.265 | 4.000 | 55.216.333 |
| Bekerja | 2631 | 1.857.642 | 4.012.856 | 7.000 | 103.195.000 |

Berdasarkan Tabel 4.2, dapat diketahui bahwa dari total keseluruhan sebanyak 1800 ibu rumah tangga yang berstatus tidak bekerja yang ada di Indonesia, rata-rata pengeluaran rumah tangga yaitu sebesar Rp 1.456.451 dan standar deviasi sebesar Rp 2.735.265 dengan pengeluaran rumah tangga minimum sebesar Rp 4.000 dan pengeluaran rumah tangga maksimum sebesar Rp 55.216.333. Sedangkan, dari total keseluruhan sebanyak 2631 ibu rumah tangga yang berstatus bekerja yang ada di Indonesia, rata-rata pengeluaran rumah tangga yaitu sebesar Rp 1.857.642 dan standar deviasi sebesar Rp 4.012.856 dengan pengeluaran rumah tangga minimum sebesar Rp 7.000 dan pengeluaran rumah tangga maksimum sebesar Rp 103.195.000. Hal ini menunjukkan bahwa pengeluaran rumah tangga untuk ibu rumah tangga yang berstatus bekerja lebih besar dibandingkan dengan pengeluaran rumah tangga untuk ibu rumah tangga yang berstatus tidak bekerja.

4.2.5 Rata-rata Gaji Wanita

Statistika deskriptif untuk karakteristik rata-rata gaji wanita yaitu variabel independen pada level 2 (Kecamatan) disajikan dengan menggunakan tabel statistika deskriptif. Hasil karakteristik dapat dilihat pada Tabel 4.3 sebagai berikut.

Tabel 4.3 Statistika Deskriptif Untuk Rata-rata Gaji Wanita Setiap Kecamatan

| Variabel | n | Rata-rata | Standar Deviasi | Minimum | Maksimum |
|-----------------------|----------|------------------|------------------------|----------------|-----------------|
| Rata-rata gaji wanita | 99 | 1.718.091 | 1.909.957 | 150.000 | 13.800.000 |

Berdasarkan Tabel 4.3, dapat diketahui bahwa rata-rata gaji wanita (data pada level Kecamatan/level 2) sebanyak 99 Kecamatan dengan nilai rata-rata yaitu sebesar Rp 1.718.091 dan standar deviasi sebesar Rp 1.909.957 dengan rata-rata gaji wanita minimum sebesar Rp 150.000 dan rata-rata gaji wanita maksimum sebesar Rp 13.800.000. Hal ini menunjukkan bahwa rata-rata gaji wanita di Indonesia yaitu antara Rp 191.866 sampai dengan Rp 3.628.048.

4.3 Uji Multikolinieritas

Pengujian multikolinieritas dilakukan untuk mengetahui korelasi antara variabel independen dengan cara melihat nilai VIF. Data yang digunakan dalam uji multikolinearitas adalah data pada Lampiran 1. Hipotesis uji multikolinearitas dapat dilihat sebagai berikut.

Hipotesis :

H_0 : Terdapat multikolinieritas antara variabel independen ($VIF > 10$)

H_1 : Tidak terdapat multikolinieritas antara variabel independen ($VIF < 10$)

Statistik uji :

$$VIF = \frac{1}{1 - R_i^2}$$

Tingkat signifikan : $\alpha = 0,05$

Daerah kritis : Tolak H_0 jika $P\text{-value} < \alpha$

Tabel 4.4 Nilai VIF Pada Variabel Independen

| Prediktor | VIF |
|------------------|------------|
| X_1 | 1,258 |
| X_2 | 1,193 |
| X_3 | 1,060 |

Berdasarkan Tabel 4.4, dapat diketahui bahwa nilai VIF dari masing-masing variabel independen kurang dari 10 maka tolak H_0 sehingga dapat disimpulkan bahwa tidak terdapat multikolinieritas antara variabel independen.

4.4 Analisis Regresi Logistik Biner

Analisis regresi logistik biner dilakukan pada data level 1 untuk mengetahui pengaruh hubungan antara variabel status bekerja ibu rumah tangga dengan variabel

pendidikan terakhir, usia dan pengeluaran rumah tangga di Indonesia, dengan data seperti pada Lampiran 1.

4.4.1 Uji Estimasi Parameter Secara Serentak

Analisis regresi logistik biner yang pertama adalah melakukan pengujian estimasi parameter secara serentak, yang digunakan untuk mengetahui pengaruh semua variabel independen secara keseluruhan terhadap status bekerja ibu rumah tangga. Hipotesis untuk pengujian estimasi parameter secara serentak dapat dilihat sebagai berikut.

Hipotesis :

$H_0: \beta_1 = \beta_2 = \beta_3 = 0$ (Variabel pendidikan terakhir ibu rumah tangga, usia ibu rumah tangga dan pengeluaran rumah tangga tidak signifikan terhadap status bekerja ibu rumah tangga).

H_1 : Minimal ada satu $\beta_i \neq 0$; $i = 1, 2, 3$ (Minimal ada salah satu dari variabel pendidikan terakhir ibu rumah tangga, usia ibu rumah tangga dan pengeluaran rumah tangga yang signifikan terhadap status bekerja ibu rumah tangga).

Statistik uji :

$$G^2 = -2 \left[\ln L(\hat{w}) - \ln L(\hat{\Omega}) \right]$$

Tingkat signifikansi : $\alpha = 0,05$

Daerah kritis : Tolak H_0 jika $G^2 > \chi^2_{(k)}$ atau P-value $< \alpha$

Tabel 4.5 Uji Estimasi Parameter Secara Serentak Analisis Regresi Logistik Biner

| | Chi-square | P-value |
|--------------|-------------------|----------------|
| Step | 108,905 | 0,000 |
| Block | 108,905 | 0,000 |
| Model | 108,905 | 0,000 |

Berdasarkan Tabel 4.5, dapat diketahui bahwa nilai $G^2 = 108,905 > \chi^2_{(0,05;6)} = 1,635$ atau P-value = $0,000 < \alpha = 0,05$, sehingga diperoleh keputusan tolak H_0 . Artinya, minimal ada salah satu dari variabel pendidikan terakhir ibu rumah tangga, usia ibu rumah tangga dan pengeluaran rumah tangga yang signifikan terhadap status bekerja ibu rumah tangga.

4.4.2 Uji Estimasi Parameter Secara Parsial

Setelah dilakukan pengujian estimasi parameter secara serentak, maka dilanjutkan dengan pengujian estimasi parameter secara parsial untuk mengetahui pengaruh dari masing-masing variabel independen terhadap status bekerja ibu rumah tangga. Hasil pengujian estimasi parameter secara parsial dapat dilihat pada Tabel 4.6 sebagai berikut.

Tabel 4.6 Uji Estimasi Parameter Secara Parsial Analisis Regresi Logistik Biner

| Variabel | Estimasi | Z | P-value |
|-------------------------|-----------------|----------|----------------|
| X₁(0) | 0,033 | 0,226 | 0,821 |
| X₁(1) | -0,148 | -0,930 | 0,352 |
| X₁(2) | -0,194 | -1,229 | 0,219 |
| X₁(3) | 0,978 | 5,070 | 0,000 |
| X₂ | 0,005 | 1,714 | 0,087 * |
| X₃ | 0,027 | 2,252 | 0,024 |
| Constant | 0,141 | 0,693 | 0,488 |

*tidak signifikan pada $\alpha = 0,05$

Berdasarkan Tabel 4.6, dapat diperoleh hipotesis pengujian estimasi parameter secara parsial untuk variabel independen yang tidak signifikan sebagai berikut.

- **Usia (X₂)**

Hipotesis:

H₀ : $\beta_2 = 0$ (Variabel usia tidak berpengaruh terhadap variabel status bekerja ibu rumah tangga)

H₁ : $\beta_2 \neq 0$ (Variabel usia berpengaruh terhadap variabel status bekerja ibu rumah tangga)

Statistik uji :

$$Z = \frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}$$

Tingkat signifikansi : $\alpha = 0,05$

Daerah kritis : Tolak H₀ jika $Z > Z_{\alpha/2}$ atau P-value $< \alpha$

Keputusan : Gagal tolak H₀ karena $Z = 1,714 < Z_{\alpha/2} = 1,96$ atau P-value = $0,087 > \alpha = 0,05$

Kesimpulan : Dapat diketahui bahwa variabel usia tidak berpengaruh terhadap variabel status bekerja ibu rumah tangga.

Dari hasil pengujian estimasi parameter untuk masing-masing variabel independen, dapat diketahui bahwa variabel usia (X_2) tidak berpengaruh terhadap status bekerja ibu rumah tangga. Sehingga, perlu dilakukan analisis regresi logistik biner tanpa mengikutsertakan variabel usia (X_2). Hasil analisis dapat dilihat pada Tabel 4.7 sebagai berikut.

Tabel 4.7 Uji Estimasi Parameter Secara Parsial Analisis Regresi Logistik Biner Tanpa X_2

| Variabel | Estimasi | Z | P-value |
|----------|----------|--------|---------|
| $X_1(0)$ | -0,026 | -0,184 | 0,854 |
| $X_1(1)$ | -0,252 | -1,716 | 0,086 |
| $X_1(2)$ | -0,300 | -2,057 | 0,040 |
| $X_1(3)$ | 0,878 | 4,777 | 0,000 |
| X_3 | 0,027 | 2,308 | 0,021 |
| Constant | 0,407 | 3,096 | 0,002 |

Berdasarkan Tabel 4.7, dapat diketahui bahwa hasil pengujian estimasi parameter untuk masing-masing variabel independen, variabel pendidikan terakhir (X_1) dan variabel pengeluaran rumah tangga (X_3) berpengaruh terhadap status bekerja ibu rumah tangga. Sehingga, model regresi logistik biner untuk status bekerja ibu rumah tangga sebagai berikut.

$$\ln\left(\frac{\pi(X)}{1-\pi(X)}\right) = 0,407 - 0,026X_1(0) - 0,252X_1(1) - 0,300X_1(2) + 0,878X_1(3) + 0,027X_3$$

Hasil analisis regresi logistik biner dari status bekerja ibu rumah tangga pada level 1 (individu) memperoleh nilai *Deviance* sebesar 5879,9.

4.5 Analisis Regresi Multilevel Respon Biner Untuk *Random Intercept*

Analisis regresi multilevel yang digunakan dalam penelitian ini adalah analisis regresi 2-level. Analisis regresi 2-level dengan respon biner yaitu untuk random intercept, yang pemodelan dilakukan tanpa mengikutsertakan variabel Z (variabel rata-rata pendapatan wanita/variabel pada level 2/Kecamatan) dan kemudian pemodelan dilakukan dengan mengikutsertakan variabel Z (variabel rata-rata pendapatan wanita/variabel pada level 2/Kecamatan), dengan data seperti pada Lampiran 2.

4.5.1 Analisis Regresi 2-Level Respon Biner Tanpa Variabel Rata-rata Gaji Wanita

Analisis regresi 2-level dengan respon biner dengan random intercept yang pertama dilakukan pemodelan tanpa mengikutsertakan variabel Z (variabel rata-rata pendapatan wanita/variabel pada level 2/Kecamatan). Hasil analisis dapat dilihat pada Tabel 4.8 sebagai berikut.

Tabel 4.8 Uji Estimasi Parameter Analisis Regresi Multilevel Tanpa Z

| Variabel | Estimasi | Z | P-value |
|-------------------------|----------|--------|---------|
| X₁(0) | 0,102 | 0,675 | 0,500 |
| X₁(1) | -0,076 | -0,456 | 0,649 |
| X₁(2) | -0,106 | -0,621 | 0,534 |
| X₁(3) | 1,103 | 5,358 | 0,000 |
| X₂ | 0,005 | 1,732 | 0,083 * |
| X₃ | 0,040 | 3,123 | 0,002 |
| Constant | 0,029 | 0,128 | 0,898 |

*tidak signifikan pada $\alpha = 0,05$

Berdasarkan Tabel 4.8, dapat diperoleh hipotesis pengujian estimasi parameter secara parsial untuk variabel independen yang tidak signifikan sebagai berikut.

- **Usia (X₂)**

Hipotesis:

H₀ : $\beta_2 = 0$ (Variabel usia tidak berpengaruh terhadap variabel status bekerja ibu rumah tangga)

H₁ : $\beta_2 \neq 0$ (Variabel usia berpengaruh terhadap variabel status bekerja ibu rumah tangga)

Statistik uji :

$$Z = \frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}$$

Tingkat signifikansi : $\alpha = 0,05$

Daerah kritis : Tolak H₀ jika $Z > Z_{\alpha/2}$ atau P-value $< \alpha$

Keputusan : Gagal tolak H₀ karena $Z = 1,732 < Z_{\alpha/2} = 1,96$ atau P-value = $0,083 > \alpha = 0,05$

Kesimpulan : Dapat diketahui bahwa variabel usia tidak berpengaruh terhadap variabel status bekerja ibu rumah tangga.

Dari hasil pengujian estimasi parameter untuk masing-masing variabel independen, dapat diketahui bahwa variabel usia (X_2) tidak berpengaruh terhadap status bekerja ibu rumah tangga. Sehingga, perlu dilakukan analisis regresi 2-level tanpa mengikutsertakan variabel usia (X_2). Hasil analisis dapat dilihat pada Tabel 4.9 sebagai berikut.

Tabel 4.9 Uji Estimasi Parameter Analisis Regresi Multilevel Tanpa Z dan Tanpa X_2

| Variabel | Estimasi | Z | P-value |
|----------|----------|--------|---------|
| $X_1(0)$ | 0.033 | 0.230 | 0.818 |
| $X_1(1)$ | -0.196 | -1.289 | 0.197 |
| $X_1(2)$ | -0.235 | -1.530 | 0.126 |
| $X_1(3)$ | 0.974 | 5.077 | 0,000 |
| X_3 | 0.040 | 3.169 | 0.002 |
| Constant | 0.330 | 2.281 | 0.023 |

Berdasarkan Tabel 4.9, dapat diketahui bahwa hasil pengujian estimasi parameter untuk masing-masing variabel independen dengan adanya pengaruh level 2 (Kecamatan), variabel pendidikan terakhir (X_1) dan variabel pengeluaran rumah tangga (X_3) berpengaruh terhadap status bekerja ibu rumah tangga. Sehingga, model regresi 2-level untuk status bekerja ibu rumah tangga sebagai berikut.

$$\ln \left(\frac{\pi(X_{ij})}{1 - \pi(X_{ij})} \right) = 0,330 + 0,033X_{li}(0) - 0,196X_{li}(1) - 0,235X_{li}(2) + 0,974X_{li}(3) + 0,040X_{3i}$$

Hasil analisis regresi 2-level dari status bekerja ibu rumah tangga memperoleh nilai *Deviance* sebesar 5792,2.

4.5.2 Analisis Regresi 2-Level Respon Biner Dengan Variabel Rata-rata Gaji Wanita

Setelah dilakukan analisis regresi 2-level dengan respon biner tanpa mengikutsertakan variabel Z (variabel rata-rata pendapatan wanita/variabel pada level 2/Kecamatan), maka dilanjutkan melakukan pemodelan dengan mengikutsertakan variabel Z (variabel rata-rata pendapatan wanita/variabel pada level 2/Kecamatan). Hasil analisis dapat dilihat pada Tabel 4.10 sebagai berikut.

Tabel 4.10 Uji Estimasi Parameter Analisis Regresi Multilevel Dengan Z

| Variabel | Estimasi | Z | P-value |
|----------|----------|-------|---------|
| $X_1(0)$ | 0.103 | 0.684 | 0.494 |

| | | | |
|-------------------------|--------|--------|-------|
| X₁(1) | -0.075 | -0.447 | 0.655 |
| X₁(2) | -0.104 | -0.608 | 0.543 |
| X₁(3) | 1.107 | 5.380 | 0,000 |
| X₂ | 0,005 | 1.766 | 0.077 |
| X₃ | 0,040 | 3.120 | 0,002 |
| Z | -0.032 | -1.031 | 0.303 |
| Constant | 0.077 | 0.333 | 0.739 |

*tidak signifikan pada $\alpha = 0,05$

Berdasarkan Tabel 4.10, dapat diperoleh hipotesis pengujian estimasi parameter secara parsial untuk variabel independen yang tidak signifikan sebagai berikut.

- **Usia (X₂)**

Hipotesis:

H₀ : $\beta_2 = 0$ (Variabel usia tidak berpengaruh terhadap variabel status bekerja ibu rumah tangga)

H₁ : $\beta_2 \neq 0$ (Variabel usia berpengaruh terhadap variabel status bekerja ibu rumah tangga)

Statistik uji :

$$Z = \frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}$$

Tingkat signifikansi : $\alpha = 0,05$

Daerah kritis : Tolak H₀ jika $Z > Z_{\alpha/2}$ atau P-value $< \alpha$

Keputusan : Gagal tolak H₀ karena $Z = 1,766 < Z_{\alpha/2} = 1,96$ atau P-value = $0,077 > \alpha = 0,05$

Kesimpulan : Dapat diketahui bahwa variabel usia tidak berpengaruh terhadap variabel status bekerja ibu rumah tangga.

- **Rata-rata Gaji Wanita (Z)**

Hipotesis:

H₀ : $\beta_{01} = 0$ (Variabel rata-rata gaji wanita tidak berpengaruh terhadap variabel status bekerja ibu rumah tangga)

H₁ : $\beta_{01} \neq 0$ (Variabel rata-rata gaji wanita berpengaruh terhadap variabel status bekerja ibu rumah tangga)

Statistik uji :

$$Z = \frac{\hat{\beta}_{01}}{SE(\hat{\beta}_{01})}$$

Tingkat signifikansi : $\alpha = 0,05$

Daerah kritis : Tolak H_0 jika $Z > Z_{\alpha/2}$ atau P-value $< \alpha$

Keputusan : Gagal tolak H_0 karena $|Z| = 1,031 < Z_{\alpha/2} = 1,96$ atau P-value = $0,303 > \alpha = 0,05$

Kesimpulan : Dapat diketahui bahwa variabel rata-rata gaji wanita tidak berpengaruh terhadap variabel status bekerja ibu rumah tangga.

Dari hasil pengujian estimasi parameter untuk masing-masing variabel independen dengan adanya pengaruh level 2 (Kecamatan), variabel pendidikan terakhir (X_1) dan variabel pengeluaran rumah tangga (X_3) berpengaruh terhadap status bekerja ibu rumah tangga. Sehingga, model regresi 2-level untuk status bekerja ibu rumah tangga sebagai berikut.

$$\ln \left(\frac{\pi(X_{ij})}{1 - \pi(X_{ij})} \right) = 0,077 + 0,103X_{1i}(0) - 0,075X_{1i}(1) - 0,104X_{1i}(2) + 1,107X_{1i}(3) + 0,040X_{3i}$$

Hasil analisis regresi 2-level dari status bekerja ibu rumah tangga memperoleh nilai *Deviance* sebesar 5788.

4.6 Pemilihan Model Terbaik

Dari hasil analisis yang telah diperoleh dalam analisis regresi logistik dan analisis regresi 2-level, dapat diketahui nilai *Deviance* dari masing-masing model yang dihasilkan. Pemilihan model terbaik dilakukan dengan cara membandingkan nilai *Deviance* setiap model. Hasil dapat dilihat pada Tabel 4.11 sebagai berikut.

Tabel 4.11 Nilai *Deviance* Setiap Model

| Model | Nilai <i>Deviance</i> |
|-------------------------------------------------------------|-----------------------|
| Regresi logistik biner | 5879,9 |
| Regresi 2-level dengan random intercept (tanpa variabel Z) | 5792,2 |
| Regresi 2-level dengan random intercept (dengan variabel Z) | 5788 |

Berdasarkan Tabel 4.11, dapat diketahui bahwa nilai *Deviance* dari model regresi 2-level dengan random intercept (tanpa variabel Z) lebih baik dibandingkan dengan nilai *Deviance* model regresi logistik biner. Sedangkan, nilai *Deviance* model regresi 2-level dengan random intercept (dengan variabel Z) lebih baik dibandingkan dengan nilai *Deviance* model regresi 2-level dengan random intercept (tanpa variabel Z), walaupun variabel independen (Z) tidak signifikan. Sehingga, model regresi 2-level dengan random intercept (dengan variabel Z) merupakan model terbaik karena nilai *Deviance* yang dihasilkan paling kecil diantara nilai *Deviance* dari model lainnya. Hasil model regresi 2-level dengan respon biner sebagai berikut.

- Model Level 1

$$\ln \left(\frac{\pi(X_{ij})}{1 - \pi(X_{ij})} \right) = \beta_{0j} + \beta_1 X_{1i}(0) + \beta_1 X_{1i}(1) + \beta_1 X_{1i}(2) + \beta_1 X_{1i}(3) + \beta_3 X_{3i}$$

- Model Level 2

$$\beta_{0j} = \beta_{00} + \beta_{01} Z_j$$

Sehingga, diperoleh model kombinasi anatara level 1 dan level 2 yang terbentuk dapat dilihat sebagai berikut.

$$\ln \left(\frac{\pi(X_{ij})}{1 - \pi(X_{ij})} \right) = (\beta_{00} + \beta_{01} Z_j) + \beta_1 X_{1i}(0) + \beta_1 X_{1i}(1) + \beta_1 X_{1i}(2) + \beta_1 X_{1i}(3) + \beta_3 X_{3i}$$

$$\ln \left(\frac{\pi(X_{ij})}{1 - \pi(X_{ij})} \right) = 0,077 + 0,103 X_{1i}(0) - 0,075 X_{1i}(1) - 0,104 X_{1i}(2) + 1,107 X_{1i}(3) + 0,040 X_{3i}$$

Dari model terbaik yang telah diperoleh, dapat diketahui bahwa ibu rumah tangga yang memiliki latar belakang tidak sekolah dan ibu rumah tangga yang memiliki latar belakang pendidikan terakhir SMA meningkatkan keputusan ibu rumah tangga untuk ikut serta dalam dunia kerja, serta semakin tingginya pengeluaran rumah tangga meningkatkan keputusan ibu rumah tangga untuk ikut serta dalam dunia kerja.

4.7 Ketepatan Klasifikasi

Ketepatan klasifikasi digunakan untuk mengetahui besarnya klasifikasi yang dihasilkan, dengan cara melihat nilai observasi dengan nilai prediksi sudah tepat atau tidak. Hasil ketepatan klasifikasi dapat dilihat pada Tabel 4.12 sebagai berikut.

Tabel 4.12 Ketepatan Klasifikasi Untuk Analisis Regresi Logistik Biner

| Aktual | | Prediksi | | Persentase Tepat |
|------------------------|---|----------|------|------------------|
| | | Y | | |
| | | 0 | 1 | |
| Y | 0 | 0 | 1800 | ,0 |
| | 1 | 0 | 2631 | 100,0 |
| Persentase Keseluruhan | | | | 59,4 |

Sehingga, dapat dihitung nilai ketepatan klasifikasi sebagai berikut.

$$1 - APER = \frac{0 + 2631}{4431} \times 100\% = 59,4\%$$

Jadi, dapat diketahui bahwa nilai ketepatan klasifikasi untuk analisis regresi logistik biner adalah 59,4% dan nilai kesalahan klasifikasi sebesar 40,6%. Nilai sensitivitas dapat dihitung sebagai berikut.

$$Sensitivitas = \frac{2631}{2631} \times 100\% = 100\%$$

Nilai sensitivitas untuk analisis regresi logistik biner adalah 100%, artinya bahwa kategori variabel Y=1 antara nilai aktual dan prediksi adalah sama. Selanjutnya, dianalisis ketepatan klasifikasi untuk analisis multilevel. Hasil ketepatan klasifikasi dapat dilihat pada Tabel 4.13 sebagai berikut.

Tabel 4.13 Ketepatan Klasifikasi Untuk Analisis Regresi Multilevel

| Tabel 10.10: Hasil Uji Akurasi Model Regresi Logistik | | | | |
|-------------------------------------------------------|---|----------|------|------------------|
| Aktual | | Prediksi | | Persentase Tepat |
| | | Y | | |
| | | 0 | 1 | |
| Y | 0 | 16 | 1784 | 0,9 |
| | 1 | 18 | 2613 | 99,3 |
| Persentase Keseluruhan | | | | 59,3 |

Sehingga, dapat dihitung nilai ketepatan klasifikasi sebagai berikut.

$$1 - APER = \frac{16 + 2613}{4431} \times 100\% = 59,3\%$$

Jadi, dapat diketahui bahwa nilai ketepatan klasifikasi untuk analisis regresi logistik biner adalah 59,3% dan nilai kesalahan klasifikasi sebesar 40,7%. Nilai sensitivitas dapat dihitung sebagai berikut.

$$Sensitivitas = \frac{2613}{2631} \times 100\% = 99,3\%$$

Nilai sensitivitas untuk analisis regresi logistik biner adalah 99,3%, artinya bahwa terdapat 0,7% dari kategori variabel $Y=1$ antara nilai aktual dan prediksi yang kurang tepat.

(Halaman ini sengaja dikosongkan)

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan, diperoleh kesimpulan sebagai berikut.

1. Estimasi parameter yang telah diperoleh untuk pemodelan menggunakan analisis regresi 2-level dengan respon biner dengan pendekatan integral kuadrat Gauss-Hermite Quadrature tidak *close form*, sehingga perlu dilakukan perhitungan iterasi Newton-Rahpson dimana nilai awal β dicari dengan menggunakan rumus GLS (*Generalized Least Square*) sebagai berikut.

$$\hat{\beta} = (X^T V^{-1} X)^{-1} X^T V^{-1} Y$$

2. Dari hasil analisis status bekerja ibu rumah tangga menggunakan analisis regresi logistik, analisis regresi 2-level dengan random intercept tanpa dan dengan mengikutsertakan variabel rata-rata gaji wanita (variabel Z), dapat disimpulkan bahwa model terbaik untuk status bekerja ibu rumah tangga adalah model regresi 2-level dengan random intercept (dengan mengikutsertakan variabel Z) lebih baik karena nilai *Deviance* paling kecil yaitu sebesar 5788. Artinya, keputusan ibu rumah tangga untuk ikut serta dalam dunia kerja juga dipengaruhi oleh adanya daerah tempat tinggal (Level 2/Kecamatan) dengan variabel rata-rata gaji tenaga kerja wanita yang berpengaruh sangat kecil. Dari model terbaik tersebut, variabel yang mempengaruhi ibu rumah tangga untuk ikut serta dalam dunia kerja adalah variabel pendidikan terakhir dan variabel pengeluaran rumah tangga. Dari hasil ketepatan klasifikasi, dapat disimpulkan bahwa nilai ketepatan klasifikasi untuk analisis regresi logistik biner dan analisis regresi multilevel dengan respon biner tidak berbeda jauh.

5.2 Saran

Berdasarkan hasil analisis dan kesimpulan yang diperoleh, saran yang diperoleh sebagai berikut.

1. Model yang digunakan dalam penelitian ini dibatasi untuk pemodelan status bekerja ibu rumah tangga menggunakan model random intercept. Pemodelan status bekerja ibu rumah tangga dengan model koefisien acak atau model dengan interaksi (*Cross-interactive model*) perlu dikembangkan untuk mengatasi permasalahan kasus multilevel yang lebih rumit yaitu dengan adanya interaksi antar variabel independen.
2. Penelitian dengan pemodelan regresi 3-level yaitu level 1 pada individu, level 2 pada tingkat Kecamatan dan level 3 pada tingkat Kabupaten/Kota perlu dikembangkan untuk mengetahui faktor-faktor yang diduga berpengaruh terhadap status bekerja ibu rumah tangga di tingkat Kabupaten/Kota.

DAFTAR PUSTAKA

- Anindita, S. D. 2016. Pemodelan Faktor-Faktor yang Mempengaruhi Ibu Rumah Tangga Untuk Bekerja Di Provinsi Papua Menggunakan Analisis Regresi Logistik Biner. Tugas Akhir. Surabaya : Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam. Institut Teknologi Sepuluh Nopember.
- Antika, S. F. 2011. Analisis Regresi Multilevel Terhadap Faktor-Faktor Yang Mempengaruhi Nilai UASBN SD/MI (Studi Kasus Nilai UASBN SD/MI di Kecamatan Tulangan Tahun Ajaran 2009/2010). Tugas Akhir. Surabaya : Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam. Institut Teknologi Sepuluh Nopember.
- Agresti. 2007. Categorical Data Analysis Second Edition. John Willey & Sons. New York.
- Alteza dan Hidayati. 2009. Work-Family Conflict Pada Wanita Bekerja : Studi Tentang Penyebab, Dampak dan Strategi Coping : Laporan Penelitian. Yogyakarta : Fakultas Ilmu Sosial dan Ekonomi Universitas Negeri Yogyakarta.
- Draper, N. dan Smith, H. 1992. Analisis Regresi Terapan edisi kedua. Jakarta : PT Gramedia Pustaka Utama.
- Dubeck, P. J. dan Borman, K. 1996. Women and Work. New York : Garland Publisher.
- Fong, Y., Rue, H. dan Wakefield, J. 2009. *Bayesian inference for generalized linear mixed models*. Biostatistics. Published by Oxford University Press.
- Ghozali, I. 2011. Aplikasi Analisis Multivariate Dengan Program SPSS. Semarang : Badan Penerbit Universitas Diponegoro.
- Goldstein, H. 1995. Multilevel Statistical Models 2nd Ed. London : E-Book of Arnold.
- Greene, W. H. 2002. Econometric Analysis, Fifth Edition. Prentice Hall. New Jersey.
- Gujarati, D. N. 2004. Basic Econometric, Fourth Edition. New York : The McGraw-Hill Companies.

- Guo, G. dan Zhao, H. 2000. *Multilevel Modeling for Binary Data*. Annual Review of Sociology, Vol 26, pp 441-462.
- Hosmer, D. W., dan Lemeshow, S. 1989. *Applied Logistic Regression*. New York: John Wiley & Sons.
- Hox, J. J. 2002. *Multilevel Analysis : Techniques and Applications*. London : Lawrence Erlbaum Associates Publishers.
- Johnson, R. A. dan Wichern, D. W. 2007. *Applied Multivariate Data Analysis*. Sixth Edition. United State : Pearson Prentice Hall, New Jersey.
- Jones, B. S. dan Steenbergen, M. 1997. *Modeling Multilevel Data Structure*. Sociological Methods and Research, Vol 22. 3, pp. 283-299.
- Kaufman, B. E. dan Hotchkiss, J. L. 2000. *The Economics of Labor Markets*. Orlando : The Dryden Press.
- Khan, M. H. R. dan Shaw, J. E. H. 2011. *Multilevel Logistic Regression Analysis Applied to Binary Contraceptive Prevalence Data*. Journal of Data Science 9, pp 93-110.
- Kutner, M.H., Nachtsheim, C.J., dan Neter, J. 2004. *Applied Linear Regression Models* 4th ed. New York : McGraw-Hill Companies, Inc.
- Lemme, B. H. 1995. *Development in Adulthood*. Boston : Allyn and Bacon Publishing.
- Linandar, T. N. 2009. *Faktor-Faktor yang Mempengaruhi Karir Wanita (Studi Kasus : Badan Pemberdayaan Masyarakat dan Keluarga Berencana Kota Bogor)*. Skripsi. Bogor : Fakultas Ekonomi dan Manajemen. Institut Teknologi Bogor.
- Majid, F. 2012. *Faktor-faktor yang Mempengaruhi Keputusan Perempuan Berstatus Menikah untuk Bekerja (Studi Kasus : Kota Semarang)*. Skripsi. Semarang : Fakultas Ekonomika dan Bisnis. Universitas Diponegoro.
- Nazir. 2010. *Analisis Determinan Pendapatan Pedagang Kaki Lima di Kabupaten Aceh Utara*. Tesis. Medan : Universitas Sumatera Utara.
- Nelson, D. L. dan Quick, J. C. 2010. *Organizational Behavior-Science The Real World and You*. Mason. USA : Cengage Learning.
- Poedjiati, S. A. 2009. *Perbandingan Regresi Logistik dan Model Multilevel Pada Data Ordinal dan Biner Untuk Mengetahui Variabel Yang Mempengaruhi*

- Hasil Belajar Mengajar Kalkulus I di ITS. Tesis. Surabaya : Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam. Institut Teknologi Sepuluh Nopember.
- Rachman, H.P.S dan Wahida. 1998. Dinamika Pola Pengeluaran Dan Konsumsi Rumah tangga Serta Prospek Permintaan Pangan dalam Dinamika Ekonomi Perdesaan: Perubahan Struktur Pendapatan, Ketenagakerjaan dan Pola Konsumsi Rumah Tangga. Bogor : Kerjasama Puslit Sosial Ekonomi Pertanian dengan Ford Foundation.
- Raudenbush, S. W. & Bryk, A. S. (2002). Hierarchical Linear Models. Applications and Data Analysis Methods (2nd ed.). Thousand Oaks. London : Sage Publications.
- Rodriguez, G. dan Goldman, N. 2001. *Improved estimation procedures for multilevel models with binary response: a case-study*. Journal Royal Statist. Soc A, 164, Part 2 pp 339-355.
- Sajogyo, P. 1983. Peranan Wanita Dalam Perkembangan Masyarakat Desa. Jakarta : CV. Rajawali.
- Simanjuntak, P. J. 1998. Pengantar Ekonomi Sumber Daya Manusia. Jakarta : LPFE-UI.
- Sinungan, M. 2003. Produktivitas : Apa dan Bagaimana. Jakarta : Bumi Aksara.

(Halaman ini sengaja dikosongkan)

Lampiran 1. Data Penelitian yang Digunakan pada Analisis Regresi Logistik Biner (Level 1)

| Kecamatan | Individu | Variabel Dependen | Variabel Independen Level 1 | | |
|-----------|----------|-------------------|-----------------------------|----------------|----------------|
| | | Y | X ₁ | X ₂ | X ₃ |
| 1 | 1 | 1 | 4 | 47 | 6462000 |
| | 2 | 1 | 1 | 46 | 842000 |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| | 56 | 1 | 3 | 40 | 2353333 |
| 2 | 1 | 1 | 3 | 57 | 9134167 |
| | 2 | 1 | 4 | 41 | 4714667 |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| | 41 | 1 | 3 | 49 | 3779667 |
| 3 | 1 | 1 | 3 | 35 | 264333 |
| | 2 | 1 | 1 | 72 | 32000 |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| | 47 | 1 | 3 | 50 | 307000 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 98 | 1 | 0 | 3 | 64 | 1102333 |
| | 2 | 1 | 4 | 38 | 18074333 |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| | 18 | 1 | 3 | 28 | 2158000 |
| 99 | 1 | 0 | 4 | 62 | 3313200 |
| | 2 | 0 | 3 | 34 | 1415333 |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| | 68 | 0 | 3 | 58 | 2389783 |

Keterangan :

Y = Status Bekerja Ibu Rumah Tangga

X₁ = Pendidikan Terakhir

X₂ = Usia

X₃ = Pengeluaran Rumah Tangga

Lampiran 2. Data Penelitian yang Digunakan pada Analisis Regresi Multilevel
(Level 2)

| Kecamatan | Individu | Variabel Dependen | Variabel Independen Level 1 | | | Variabel Independen Level 1 |
|-----------|----------|----------------------|-----------------------------|----------------|----------------|-----------------------------------|
| | | Y | X ₁ | X ₂ | X ₃ | Z |
| 1 | 1 | 1 | 4 | 47 | 6462000 | 1200000 |
| | 2 | 1 | 1 | 46 | 842000 | 1200000 |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| | 56 | 1 | 3 | 40 | 2353333 | 1200000 |
| 2 | 1 | 1 | 3 | 57 | 9134167 | 2250000 |
| | 2 | 1 | 4 | 41 | 4714667 | 2250000 |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| | 41 | 1 | 3 | 49 | 3779667 | 2250000 |
| 3 | 1 | 1 | 3 | 35 | 264333 | 1500000 |
| | 2 | 1 | 1 | 72 | 32000 | 1500000 |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| | 47 | 1 | 3 | 50 | 307000 | 1500000 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 98 | 1 | 0 | 3 | 64 | 1102333 | 1500000 |
| | 2 | 1 | 4 | 38 | 18074333 | 1500000 |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| | 18 | 1 | 3 | 28 | 2158000 | 1500000 |
| 99 | 1 | 0 | 4 | 62 | 3313200 | 2100000 |
| | 2 | 0 | 3 | 34 | 1415333 | 2100000 |
| | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| | 68 | 0 | 3 | 58 | 2389783 | 2100000 |

Keterangan :

Y = Status Bekerja Ibu Rumah Tangga

X₁ = Pendidikan Terakhir

X₂ = Usia

X₃ = Pengeluaran Rumah Tangga

Z = Rata-rata Gaji Tenaga Kerja Wanita

Lampiran 3. Syntax *Software R* Untuk Analisis Regresi Logistik Biner

```
y <- (data$Y, levels=0:1)
x1 <- factor(data$X1, levels=0:4)
x2 <- data$X2
x3 <- data$X3

Y <- (y, ref="1")
X1 <- (x1, ref="4")

logit <- glm(y ~ x1 + x2 + x3, data = data, family = binomial(link='logit'))

summary(logit)

backward<-step(logit, direction='backward')

wald.test(b=coef(logit), Sigma=vcov(logit),Term=1)
```

Lampiran 4. Syntax *Software R* Untuk Analisis Regresi 2-Level Tanpa Variabel Z

```
library(lme4)

data <- within(data,{
  y <- (data$Y)
  x1 <- (data$X1)
  x2 <- data$X2
  x3 <- data$X3
})

multilevel <- glmer(y ~ x1 + x2 + x3, data = data, family = binomial, control =
glmerControl(optimizer = "bobyqa"),
  nAGQ = 0)

summary(multilevel)
```

Lampiran 5. Syntax *Software R* Untuk Analisis Regresi 2-Level Dengan Variabel Z

```
library(lme4)

data <- within(data,{
  y <- (data$Y)
  x1 <- (data$X1)
  x2 <- data$X2
  x3 <- data$X3
  z <- data$Z
})

multilevel <- glmer(y ~ x1 + x2 + x3 + z, data = data, family = binomial, control =
glmerControl(optimizer = "bobyqa"),
  nAGQ = 0)

summary(multilevel)
```

Lampiran 6. Output *Software R* Untuk Analisis Regresi Logistik Biner

```
Call:
glm(formula = y ~ x1 + x2 + x3, family = binomial(link = "logit"),
    data = data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.3466 -1.2675  0.9726  1.0449  1.1632

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.140970  0.203340  0.693  0.4881
x11          0.032638  0.144262  0.226  0.8210
x12         -0.147773  0.158842 -0.930  0.3522
x13         -0.194261  0.158124 -1.229  0.2192
x14          0.977880  0.192885  5.070 3.98e-07 ***
x2           0.004537  0.002647  1.714  0.0865 .
x3           0.026669  0.011842  2.252  0.0243 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 5985.9 on 4430 degrees of freedom
Residual deviance: 5877.0 on 4424 degrees of freedom
AIC: 5891

Number of Fisher Scoring iterations: 4

Start: AIC=5890.99
y ~ x1 + x2 + x3

      Df Deviance  AIC
<none>    5877.0 5891.0
- x2    1  5879.9 5891.9
- x3    1  5882.8 5894.8
- x1    4  5963.3 5969.3
> #uji wald
> wald.test(b=coef(logit), Sigma=vcov(logit),Term=1)
Wald test:
-----

Chi-squared test:
X2 = 0.48, df = 1, P(> X2) = 0.49
```

Lampiran 7. Output *Software R* Untuk Analisis Regresi 2-Level Tanpa Variabel Z

Generalized linear mixed model fit by maximum likelihood (Adaptive Gauss-Hermite

Quadrature, nAGQ = 0) [glmerMod]

Family: binomial (logit)

Formula: $y \sim x1 + x2 + x3 + (1 | id)$

Data: data

Control: glmerControl(optimizer = "bobyqa")

| AIC | BIC | logLik | deviance | df.resid |
|--------|--------|---------|----------|----------|
| 5805.1 | 5856.3 | -2894.6 | 5789.1 | 4423 |

Scaled residuals:

| Min | 1Q | Median | 3Q | Max |
|---------|---------|--------|--------|--------|
| -5.4948 | -1.0383 | 0.6018 | 0.8345 | 1.4677 |

Random effects:

| Groups | Name | Variance | Std.Dev. |
|--------|------|----------|----------|
|--------|------|----------|----------|

| | | | |
|----|-------------|--------|--------|
| id | (Intercept) | 0.1876 | 0.4331 |
|----|-------------|--------|--------|

Number of obs: 4431, groups: id, 99

Fixed effects:

| | Estimate | Std. Error | z value | Pr(> z) |
|-------------|-----------|------------|---------|--------------|
| (Intercept) | 0.029020 | 0.226053 | 0.128 | 0.89785 |
| x11 | 0.101591 | 0.150456 | 0.675 | 0.49953 |
| x12 | -0.076309 | 0.167414 | -0.456 | 0.64853 |
| x13 | -0.106026 | 0.170613 | -0.621 | 0.53431 |
| x14 | 1.102548 | 0.205776 | 5.358 | 8.41e-08 *** |
| x2 | 0.004909 | 0.002834 | 1.732 | 0.08328 . |
| x3 | 0.039672 | 0.012704 | 3.123 | 0.00179 ** |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:

| (Intr) | x11 | x12 | x13 | x14 | x2 | |
|--------|--------|--------|--------|--------|--------|--------|
| x11 | -0.747 | | | | | |
| x12 | -0.810 | 0.846 | | | | |
| x13 | -0.826 | 0.847 | 0.847 | | | |
| x14 | -0.685 | 0.703 | 0.706 | 0.727 | | |
| x2 | -0.768 | 0.261 | 0.413 | 0.435 | 0.362 | |
| x3 | -0.017 | -0.043 | -0.067 | -0.113 | -0.169 | -0.030 |

Lampiran 8. Output *Software R* Untuk Analisis Regresi 2-Level Dengan Variabel Z

Generalized linear mixed model fit by maximum likelihood (Adaptive Gauss-Hermite

Quadrature, nAGQ = 0) [glmerMod]

Family: binomial (logit)

Formula: $y \sim x1 + x2 + x3 + z + (1 | id)$

Data: data

Control: glmerControl(optimizer = "bobyqa")

| AIC | BIC | logLik | deviance | df.resid |
|--------|--------|---------|----------|----------|
| 5806.0 | 5863.6 | -2894.0 | 5788.0 | 4422 |

Scaled residuals:

| Min | 1Q | Median | 3Q | Max |
|---------|---------|--------|--------|--------|
| -5.5115 | -1.0372 | 0.6010 | 0.8363 | 1.4655 |

Random effects:

| Groups Name | Variance | Std.Dev. |
|-------------|----------|----------|
|-------------|----------|----------|

| | | |
|----------------|--------|--------|
| id (Intercept) | 0.1848 | 0.4299 |
|----------------|--------|--------|

Number of obs: 4431, groups: id, 99

Fixed effects:

| | Estimate | Std. Error | z value | Pr(> z) |
|-------------|-----------|------------|---------|--------------|
| (Intercept) | 0.076728 | 0.230507 | 0.333 | 0.73923 |
| x11 | 0.102949 | 0.150473 | 0.684 | 0.49387 |
| x12 | -0.074806 | 0.167420 | -0.447 | 0.65501 |
| x13 | -0.103719 | 0.170609 | -0.608 | 0.54323 |
| x14 | 1.107463 | 0.205846 | 5.380 | 7.45e-08 *** |
| x2 | 0.005008 | 0.002835 | 1.766 | 0.07733 . |
| x3 | 0.039663 | 0.012713 | 3.120 | 0.00181 ** |
| z | -0.032443 | 0.031467 | -1.031 | 0.30253 |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:

| (Intr) | x11 | x12 | x13 | x14 | x2 | x3 | |
|--------|--------|--------|--------|--------|--------|--------|--------|
| x11 | -0.730 | | | | | | |
| x12 | -0.792 | 0.846 | | | | | |
| x13 | -0.807 | 0.847 | 0.847 | | | | |
| x14 | -0.665 | 0.703 | 0.706 | 0.727 | | | |
| x2 | -0.745 | 0.261 | 0.413 | 0.435 | 0.362 | | |
| x3 | -0.016 | -0.043 | -0.067 | -0.113 | -0.170 | -0.030 | |
| z | -0.198 | -0.011 | -0.011 | -0.015 | -0.028 | -0.035 | -0.005 |

BIODATA PENULIS



Penulis mempunyai nama lengkap Fanny Ayu Octaviana yang biasa dipanggil Fanny atau *Miss Purple*. Penulis merupakan anak kedua dari tiga bersaudara. Dilahirkan di Kota Mojokerto pada tanggal 25 Oktober 1991. Penulis telah menyelesaikan pendidikan formal yaitu TK Sacharosa Krian, SDN 4 Krian, SMPN 2 Krian, SMAK “Untung Suropati” Krian. Setelah lulus dari SMA, penulis mengikuti seleksi penerimaan mahasiswa baru dan diterima di Jurusan

Statistika Institut Teknologi Sepuluh Nopember (ITS) Surabaya untuk Program Studi Diploma 3 pada tahun 2009. Setelah lulus dari program studi Diploma, penulis langsung melanjutkan ke Program Studi Strata 1 (Lintas Jalur) Jurusan Statistika ITS pada tahun 2012. Setelah lulus dari program studi S1, penulis mendapatkan kesempatan untuk melanjutkan pendidikan pascasarjana Program Studi Magister (Strata 2) di Jurusan Statistika ITS pada tahun 2015. Penulis banyak mengucapkan terima kasih kepada dosen pembimbing, atas bantuan dan saran hingga bisa menyelesaikan sampai Program Studi Magister Statistika pada tahun 2017. Segala saran dan kritik yang membangun, sangat penulis harapkan untuk kebaikan kedepannya. Penulis dapat dihubungi melalui email : fannyayuoct@gmail.com.